NASA Conference Publication 2207

Optical Information Processing for Aerospace Applications

Proceedings of a NASA conference
held at Langley Research Center
Hampton, Virginia
August 18-19, 1981

NASA
National Aeronautics
and Space Administration

**Scientific and Technical
Information Branch**

1981

PREFACE

The Conference on Optical Information Processing for Aerospace Applications was held on August 18-19, 1981, at Langley Research Center, Hampton, Virginia. The purpose of this conference was to review current research in optical processing and to determine its role in future aerospace systems. It is clear from these proceedings that optical processing offers significant potential in the fields of aircraft and spacecraft control, pattern recognition, and robotics. The development of optical devices and components has demonstrated that system concepts can be implemented in practical aerospace configurations.

This conference was sponsored by the Office of Aeronautics and Space Technology (OAST) of NASA Headquarters and managed by Langley Research Center. Special thanks are offered to Dr. Martin Sokoloski of OAST, Rudolfo Segura (technical chairman), and the many others who gave of their time and talents.

Use of trade names or names of manufacturers in this report does not constitute an official endorsement of such products or manufacturers, either expressed or implied, by the National Aeronautics and Space Administration.

Robert L. Stermer
Chairman

CONTENTS

PREFACE iii

SESSION I - INFORMATION PROCESSING REQUIREMENTS FOR AEROSPACE APPLICATIONS

Chairman: Norman J. Berg

1. AIR FORCE RESEARCH IN OPTICAL PROCESSING 1
John Neff

2. EARTH FEATURE IDENTIFICATION FOR ONBOARD MULTISPECTRAL DATA EDITING: COMPUTATIONAL EXPERIMENTS 17
R. M. Aherron, R. F. Arduini, R. E. Davis, F. O. Huck, and S. K. Park

3. ROLE OF OPTICAL COMPUTERS IN AERONAUTICAL CONTROL APPLICATIONS 33
Robert J. Baumbick

4. OPTICAL DIGITAL COMPUTERS (Abstract only) 45
Alan Huang

5. HIGH SPEED TECHNIQUES FOR SYNTHETIC APERTURE RADAR IMAGE FORMATION (Abstract only) 47
Demetri Psaltis, Armand R. Tanguay, Jr., and Thomas J. Bicknell

SESSION II - OPTICAL SYSTEMS FOR PROCESSING INFORMATION - I

Chairman: David Casasent

6. OPTICAL IMPLEMENTATION OF SYSTOLIC ARRAY PROCESSING 53
H. J. Caulfield, W. T. Rhodes, M. J. Foster, and Sam Horvitz

7. MATRIX VECTOR MULTIPLIER WITH TIME VARYING SINGLE DIMENSIONAL SPATIAL LIGHT MODULATORS 61
A. Tarasevich, N. Zepkin, and W. T. Rhodes

8. INCOHERENT OPTICAL MATRIX-MATRIX MULTIPLIER 71
A. R. Dias

9. BULK AND INTEGRATED ACOUSTO-OPTIC SPECTROMETERS FOR RADIO ASTRONOMY 85
G. Chin, D. Buhl, and J. M. Florez

10. ENERGY NORMALIZATION OF TV VIEWED OPTICAL CORRELATION 93
Alex Grumet

SESSION III - OPTICAL SYSTEMS FOR PROCESSING INFORMATION - II

Chairman: J. L. Spencer

11. ITERATIVE OPTICAL VECTOR-MATRIX PROCESSORS 105
David Casasent and Charles Neuman

12. INTERFEROMETRIC SURFACE-WAVE ACOUSTO-OPTIC TIME-INTEGRATING CORRELATORS 119
Norman J. Berg, Irwin J. Abramovitz, and Michael W. Casseday

13. NONCOHERENT OPTICAL PROCESSING OF RASTER FORMATTED SIGNALS (Abstract only) 131
William J. Miceli and William W. Stoner

14. THREE METHODS FOR PERFORMING HANKEL TRANSFORMS 133
R. A. Athale, H. H. Szu, and J. N. Lee

15. OPTICAL RECOGNITION OF STATISTICAL PATTERNS 145
Sing H. Lee

16. OPTICAL SPATIAL INTEGRATION METHO[illegible]UITY FUNCTION GENERATION 167
P. N. Tamura, J. J. Rebholz, O. [illegible] and T. C. Lee

17. THE INTEGRATED OPTIC RF SPECTRUM ANALYZER 173
M. E. Pedinoff, T. R. Ranganath, T. R. Joseph, and J. Y. Lee

SESSION IV - INTEGRATED OPTICS AND OPTICAL COMPONENTS

Chairman: Gordon Chin

18. APPLICATIONS OF ELECTRO-OPTIC GRATINGS IN INTEGRATED OPTICAL SIGNAL PROCESSING DEVICES 209
Carl M. Verber

19. FOUR- AND FIVE-LAYER SILICON-CLAD DIELECTRIC WAVEGUIDES 217
Glen McWright and T. E. Batchman

20. CO_2 LASER WAVEGUIDING IN PROTON IMPLANTED GaAs 231
H. A. Jenkinson and D. C. Larson

21. AlGaAs/GaAs TRANSVERSE JUNCTION STRIPE LASERS WITH DISTRIBUTED FEEDBACK 241
M. J. Hafich, R. A. Skogman, P. E. Peterson, and H. Kawanishi

22. EVAPORATED As_2S_3 LUNEBURG LENSES FOR $LiNbO_3$:Ti OPTICAL WAVEGUIDES 251
James R. Busch, Van E. Wood, Richard P. Kenan, and Carl M. Verber

SESSION V - OPTICAL INFORMATION CIRCUITS

Chairman: Stephen T. Kowel

23. CURRENT AND FUTURE OPTICAL INTEGRATED CIRCUIT TRENDS FOR INFORMATION PROCESSING (Abstract only) 263
M. K. Barnoski

24. A BROADBAND RF CONTINUOUSLY VARIABLE TIME DELAY DEVICE 265
Frederick W. Freyre

25. MONOLITHICALLY INTEGRATED ACTIVE OPTICAL DEVICES 275
J. Ballantyne, D. K. Wagner, B. Kushner, and S. Wojtczuk

26. MICROCHANNEL SPATIAL LIGHT MODULATOR 291
Cardinal Warde

27. OPTICAL DATA LINKS USING LONG WAVELENGTH TECHNOLOGY (Abstract only) 299
D. J. Channin

28. INFRARED FIBER OPTIC FOCAL PLANE DISPERSERS 301
John H. Goebel

SESSION VI - OPTICAL COMPONENTS

Chairman: Herbert Hendricks

29. E-BEAM GENERATED HOLOGRAPHIC MASKS FOR OPTICAL VECTOR-MATRIX MULTIPLICATION 309
Steven M. Arnold and Steven K. Case

30. MULTI-FACET HOLOGRAPHIC OPTICAL ELEMENTS 319
Steven K. Case and Paul R. Haugen

31. SEMICONDUCTOR LASER AND DETECTOR BANDWIDTH LIMITATIONS IN OPTICAL INFORMATION PROCESSING (Abstract only) 327
Sverre T. Eng

32. A LIQUID CRYSTAL ADAPTIVE LENS 329
Stephen T. Kowel and Dennis S. Cleverly

33. MANIPULATION OF LIGHT WITH MAGNETO-OPTIC STRIPE DOMAIN FILMS 341
E. J. Torok, F. G. Hewitt, J. A. Krawczak, and G. L. Nelson

34. EFFICIENT BUTT COUPLING OF HIGH POWER CDH-LOC LASERS TO Ti-$LiNbO_3$ (LNT) OPTICAL WAVEGUIDES (Abstract only) 351
J. M. Hammer, D. Botez, C. C. Neil, and J. C. Connolly

ATTENDEES 353

AIR FORCE RESEARCH IN OPTICAL PROCESSING

John Neff
AFOSR/NE
Bolling AFB
Washington, DC 20332

SUMMARY

The Air Force is actively supporting scientific research in optical processing through the awarding of contracts and grants to university and industrial research laboratories. The emphasis is on optical and optical-electronic hybrid processing especially in the application area of image processing. A special interest exists in real-time pattern recognition processors for such airborne missions as target recognition, tracking, and terminal guidance. This paper will describe the areas of interest and the ongoing efforts in the Air Force research program.

INTRODUCTION

The history of computer technology indicates quite clearly that with the development of every new computer of larger capacity and faster speed the number of problem solving areas which would be benefitted increases and multiplies. During the past twenty years, the progress in solid state electronic computers has been extremely rapid, to the point that computers influence almost every aspect of our lives. Even though continuing progress in electronic computer technology can be anticipated, we would be remiss to overlook other techniques which show promise of complementing or even replacing the presently prevalent silicon LSI wafer. One promising alternative is provided by optics, including electro-optics and acousto-optics. And certainly not to be overlooked is the area of hybrid optical-analog/electronic-digital processing. Optical systems offer parallel computations with speeds unattainable with present digital computer technology, and the hybrid systems combine this speed and data handling capability with the computational flexibility and accuracy of digital processing. The development of efficient hybrid processors along with continued advancements in pattern recognition and computer-aided decision making will provide us with critical technology needed to maintain our world leadership in computing and signal processing capability.

One of the most promising application areas for this emerging technology is image processing. Imaging sensors are becoming an increasingly important part of such military systems as reconnaisance, missile guidance, target identification and tracking, and space surveillance. Sensors of the future are expected to provide high resolution images at frame rates which will produce data in excess of

10^{11} bits per second--a rate that digital technology will be hard pressed to handle even under the most optimistic projections. A major goal of the Air Force research in optical processing is to increase the flexibility of these processors to the extent that they can perform the functions of most of the image processing algorithms which have pushed digital technology to its limit.

I will discuss the Air Force research interests in terms of the following five categories: pattern recognition, image encoding, incoherent image processing, integrated optics and acousto-optics, and hybrid opto-electronic processing. The first three categories will receive the majority of attention since they deal with image processing.

PATTERN RECOGNITION

The basic type of optical pattern recognition system is the frequency plane correlator or matched filter shown in figure 1. The reference or target information is stored in the processor as the complex conjugate of the Fourier transform. If the target information is present in the input, a peak of light will appear in the output plane, and the location of the peak in this plane will correspond to the relative location of the target in the input scene. Figure 2 illustrates the optical correlator in a more operational scenario in which a real-time input device is employed in the input plane and a large array of potential target filters are stored in the frequency plane. The input device, known as a real-time spatial light modulator, functions as an incoherent-to-coherent image transducer. One such modulator concept currently being investigated (ref 1) is shown in figure 3, and consists of a photocathode, a micro-channel array plate, and an electro-optic plate. The electro-optic plate carries a high resistivity dielectric mirror on one side and a transparent conducting electrode on the other. The input signal illuminates the photocathode which produces an electron image of the optical intensity. The image is then amplified by the microchannel plate and deposited on the dielectric mirror. This charge distribution establishes a corresponding spatially varying refractive index in the electro-optic crystal which is capable of modulating a readout laser beam. Successful operation of this device has been achieved, and current research efforts are aimed at improving its operation so as to meet the original performance goals of 20 line pairs/millimeter spatial resolution, 1 KHz framing rate, an optical quality of less than one tenth wavelength distortion of the wavefront over a 25 millimeter diameter, a sensitivity of one nanojoule per square centimeter, and a long-term optical information storage capability on the order of weeks. Using a new internal processing mode, based on secondary electron emission, the research is proceeding to show that the device is capable of performing level thresholding, contrast reversal, contrast enchancement, image addition, and image subtraction.

A high-accuracy terminal guidance sensor system will consist of an acquisition subsystem that scans the scene in the target area, a reference subsystem containing a photographic or equivalent image of the target and its surround, and a processing subsystem that correlates the real-time scene with the reference image and generates error signals to vector the warhead onto the target. The fundamental difficulty in achieving a practical terminal guidance system lies in correlation of the reference image with a real-time image that differs in scale, aspect, contrast, and even content when sensed in a different spectral band or at a different time than the reference image. In addition to reliability and accuracy of correlation, the processor must be rapid to permit a reasonable guidance error signal bandwidth for fast target approaches. The resulting requirement is for a processing subsystem with throughput rates in excess of 10^7 operations per second, and optical correlation with its vast parallelness at low power and cost is a most promising technology.

Various optical processing approaches to missile guidance are being compared on the basis of input resolution, grey scale requirements, quantity of stored reference imagery needed, cycle times required, accuracy expected, processing needed, etc. The major emphasis has been on finding a processor or correlator architecture that is capable of maintaining correlation in the presence of various image degradations that invariably occur between the input and reference function during terminal guidance. Among the differences that have been considered are: scale, rotation, resolution, contrast, relief, data dropout, aspect, seasonal, years apart, and headings. Techniques are being studied for optimizing matched spatial filter synthesis parameters to improve the signal-to-noise ratio; however, these techniques often accentuate the high frequency detail of the image, thus rendering the correlation more susceptible to image differenes between input and reference. Therefore, space variant processing techniques are also being investigated to realize optimum correlation systems invariant to the dominant degradations. For example, the realization of a rotation-invariant correlation system can be obtained by performing a polar transform on the input and reference imagery. The magnitude of the correlation peak remains unchanged while the location of the peak translates proportional to the rotational difference between the two images. It has been shown that a multiple-invariant, space-variant optical processor is possible in which two functions described by any number of separate distortion parameters can be correlated with no loss in the signal-to-noise ratio of the correlation (ref 2). Other correlation procedures are also being investigated such as optical homomorphic filtering, optical statistical pattern recognition, an optical analog of the invariant moment method, and a hyperspace expansion and clustering technique.

Besides research into improved light modulation and novel filtering schemes for pattern recognition systems, the Air Force has an interest in holographic optical element (HOE) research. HOEs offer advantages of a reduction in the mechanical stability problems of conventional optical systems, a reduction in the number of optical elements, and a reduction in weight, size, and cost, all of which render these elements very applicable to optical systems for aircraft and missiles. The current state of the art in HOEs is limited to transfer function representation of linear, space-invariant optical systems. Unfortunately, such systems represent just a small fraction of the optical systems of interest. For example, only imaging systems with unit magnification are strictly space-invariant. For a space-variant system the impulse response is a function of the position of the impulse or point source of light in the input plane of the system. This means that in order to completely specify the impulse response, the response must be known for impulses located at all possible points in the input plane. Investigations are curently pursuing a sampling theorem-based technique for characterizing 2-D space-variant optical systems. The holograms representing the various impulse responses are multiplexed using a technique which encodes the phase of each reference beam with various diffusers. Another approach being studied involves computer multiplexing of the individual holograms. The research goals include not only a clear understanding of the basic theoretical and practical limitations of these approaches but also the discovery of additional multiplexing techniques (ref 3).

The final area of research interest under pattern recognition deals with the nonlinear processing of imagery. This opens up a whole new realm of possibilities to optical processing. For example, a basic tool in pattern recognition is the histogram which may be synthesized from a succession of intensity level slices such as shown in figure 4. Ideally, a level slice operation converts a continuous tone image into a pattern which represents only those points or areas of the input image which have a given intensity level. Figure 4 shows two non-ideal slices but ones which are sufficient to construct a histogram. The histogram would be constructed from the level slices for all possible input intensity intervals. Note, however, that the transfer function (I_{out} versus I_{in}) is a nonlinear operation. Three different approaches to nonlinear optical processing are currently being pursued: pulse-width modulation employing halftone hard-clipping and coherent filtering, direct use of light modulator nonlinear characteristics, and conversion of input intensities to spatial frequencies by employing the variable grating mode of liquid crystals. In the first approach, an input signal or image with continuous levels is converted into a halftone image that is binary. This halftone image consists of a periodic array of black dots whose sizes vary spatially according to the spatial variation of intensity in the input. This is accomplished by optically adding a halftone screen and the input image and then hard-clipping the resulting image. Nonmonotonic nonlinearities are obtained by employing various Fourier plane filters with an appropriate halftone screen design. The specific nonlinearity will depend on the halftone dot shape and on which diffraction order is passed by the filter.

The second approach is directed toward achieving nonlinearities by relying directly on nonlinear characteristics of electro-optical devices such as real-time spatial light modulators. For example, certain of these devices depend on an electro-optically controlled birefringence to produce a selective linear differential phase retardation along two axes of a crystal. Upon placing the crystal between crossed polarizers, a sinusoidal variation of intensity transmittance with voltage is obtained which can be the basis for many optically controlled nonlinear functions. One such function which has been demonstrated is three-bit optical parallel A/D conversion with a bit rate potential estimated at 1.2×10^8 points processed per second (ref 4).

The third approach, the variable grating mode technique, is based on the fact that under certain conditions a linear, phase grating structure can be established in a liquid crystal, and the period of this grating can be changed by varying an externally applied voltage as illustrated by figure 5. By employing a photoconductor in connection with the liquid crystal, it should be possible to have a grating frequency whose spatial variations correspond to the intensity variations of an image projected onto the photoconductor (ref 5). If the resulting encoded image is input to an optical spatial filtering system (figure 6), the various spatial frequency components would be diffracted to different locations in the frequency plane and could be filtered to produce an output with any desired nonlinear relationship to the input image intensities.

IMAGE ENCODING

With an increased interest in image type operations on behalf of the military comes an increased interest in image transmission. Remotely piloted vehicles and battlefield surveillance are two military operations where image communications will likely play a prominent role. Two definite problems encountered in any long distance transmission of analog image information are limited bandwidth and atmospheric scattering. With regard to the bandwidth problem, real-time image compression schemes are needed. The most severe limitation on existing schemes is the extensive computational burden, a problem for which optics may indeed be able to provide a solution. Research has successfully formulated a modified differential pulse code modulation (DPCM) technique which is suitable for incoherent optical implementation. Referred to as interpolated DPCM, or IDPCM, the technique employs a mask to spatially sample the image. These samples are then quantized and transmitted. In addition, the sampled image is interpolated by a low-pass convolution and a different image is generated by subtracting this interpolated image from the original. This difference image is also sampled, quantized, and transmitted. At the receiver, the image samples are interpolated to form a low-frequency version, and the difference-image samples are added to the interpolator output to reconstruct the original. A simulation of this technique has shown the validity of employing optical processing for bandwidth compression. Efforts are continuing toward expanding optical compression schemes into the realms of interframe and adaptive compression (ref 6).

Another approach being pursued is to employ optical transformation of the image followed by a delta modulation of the transformation. Two analog transform encoding schemes are being considered: one is a coherent optical Fourier transformation while the other is an incoherent optical Hadamard transformation. Real-time linear and adaptive delta modulators operate on the video signal obtained via a TV camera in the output plane of the optical system (ref 7). With regard to the Hadamard system, the image is broken up into a large number of sub-images by a fly's-eye lens, and each sub-image is weighted by a Hadamard mask.

The second problem area, that of atmospheric scattering, may be solvable by employing some novel encoding process such as intensity-to-frequency conversion. This involves using a wavelength-coded source to illuminate the object so that each resolution element (pixel) is illuminated with a different wavelength, thus realizing a spatial-to-spectral conversion of the object information. At the receiver, the collected light is decoded by a spectroscope which displays the various wavelengths at the proper spatial positions so as to reconstruct the original image. Since scattered light does not undergo any wavelength shift, the information on each of the spectral carriers is free to travel any path, either direct or multiply scattered, between the transmitter and the receiver. It may even be possible to use such a technique to transmit three-dimensional images (ref 8).

INCOHERENT IMAGE PROCESSING

Since the discovery of the laser in the early 1960's, coherent optical processing has advanced steadily toward developing an information processing capability with speed-of-light performance. Two major difficulties of coherent processing which have prevented a more rapid advancement are the nonavailability of suitable spatial light modulators and a relatively low dynamic range. Incoherent processing, of course, does not need the incoherent-to-coherent conversion step which the spatial light modulators perform for coherent processors. Furthermore, incoherent processing avoids the problem of coherent noise or speckle which often masks the effects of the desired image processing operation. By lessening the coherence of the system, one is, in effect, increasing the number of information bearing channels, all of which are carrying the same information in parallel. The major result of this redundancy is an increase in the signal to noise ratio. On the other hand, incoherent processing is not without problems, such as a built-up of dc-bias.

Overcoming the problems of incoherent processing, or at least significantly reducing their impact, is the goal of Air Force research in this area. If the fundamental problems can be solved, the ultimate objective will be to realize all optical processing operations with incoherent systems. Not only do incoherent processors offer a significant reduction in cost over coherent systems, but the mechanical alignment problems are much less severe, and this may be a very important point if optical processors are ever to find their way into airborne information systems.

Optical processors use interference to process information; however, with incoherent light, interference is much more difficult to achieve. Consider the formation of a Fourier transform hologram. If the required reference beam does not pass through the Fourier transforming lens, its optical path will not exactly match the object beam path and there will be no interference. But if the reference beam does pass through the lens, it will be Fourier transformed into a point image and will not be able to uniformly cover the object beam. This problem is being investigated using a technique involving a chirp-z transform. This method achieves a Fourier transform in a way basically different from the usual way, which is to place the object on the input side of a lens and take the output from the back focal plane. In the chirp-z method, a zone plate lens overlays the object, producing a compound object. The light then passes through a spatial filtering system consisting of two lenses and a spatial filter, and the resulting output is a Fourier transform. By making the Fourier transform in this manner, the previously noted problem can be avoided and it becomes possible to insert the entire optical system into a three-grating achromatic interferometer. Thus, after achromatizing the Fourier transformation process and introducing, by means of the interferometer, an achromatic coherent reference beam, it becomes possible to produce Fourier transform holograms in completely white light (ref 9).

Another effort is working more directly on conceiving an incoherent matched filter correlator (ref 10). The concept demonstrated to date employs a color-compensating grating which collects correlation-signal energy over a band of wavelengths, which has been dispersed by the matched filter, and concentrates it in a single spot to achieve a much improved S/N for the correlation peak. However, correlator operation still suffers from the very narrow spectral response of the matched filter itself. Various combinations of zone plates and gratings will be tried in an effort to overcome this deficiency and create achromatic systems capable of performing the operations of Fourier optics in white light.

A third approach to incoherent optical processing, while not as universal or flexible as the previous two techniques, has the advantage of being relatively simple to implement. A diffraction grating is employed in the input plane of the optical processor to effect a sinusoidal modulation of the input signal. This results in two off-axis spectral distributions in the Fourier plane of the processor, and these distributions will be dispersed into rainbow color due to the use of incoherent light. The origins of the differently colored signal spectra will be linearly dispersed in a direction perpendicular to the grating. A complex spatial filter for each of a discrete set of wavelengths will be placed in the spatial-frequency plane to produce a set of mutually incoherent filtered signals. The output of this incoherent filtering scheme will be formed by the incoherent addition of the discrete spectral bands (ref 11).

INTEGRATED OPTICS AND ACOUSTO-OPTICS

Spurred on by the tremendous impact that electronic integrated circuits have had, the field of integrated optics has already progressed to the point where complete circuits are being attempted. Such basic elements as optical modulators, coherent light sources, lenses with near diffraction-limited performance, and optical detectors have been demonstrated and may now be combined on a single substrate to perform some practical signal processing operations. In an effort to further expand these processing operations, the Air Force is interested in supporting research that will advance the current state-of-the-art, especially in achieving more compatibility between electronic and optical integrated systems. Much research is still needed in the areas of programmable filters, more simplistic lenses, broader band surface acoustic waves for acousto-optic processors, and higher processing gains.

Investigations are being conducted into Fresnel diffraction in an attempt to fabricate efficient and inexpensive waveguide lenses for integrated optical systems. Current lenses are either geodesic or Luneburg; however, geodesic lenses involve an expensive and time-consuming precision grinding process while Luneberg lenses have very limited focusing capabilities in high index-of-refraction waveguides because of the small change of the effective index that can be obtained from known materials. Fresnel lenses have the potential to outperform both the geodesic and Luneburg lenses. They consist of pads deposited on the waveguide which act to achieve either phase shift or absorption in alternate half-period zones (ref 12).

A recent accomplishment in the area of programmable filtering was the demonstration of a 32-bit, 17 Mbit/sec digital correlator as shown in figure 7. The reference signal is programmed into an array of interdigital electrode sets fabricated on a thin buffer layer on the surface of a planar electro-optic waveguide. The input signal is encoded as one of two possible frequencies of a surface acoustic wave (SAW). If a perfect match exists between the input and reference signals, the optical energy which is first diffracted by the SAW is diffracted a second time by the electro-optic grating produced by the interdigital electrode sets and is focused onto the detector. If a match does not exist, such as the case in the figure, some of the energy is diffracted to another point in space and does not reach the detector (ref 13).

There will always be some loss and degradation due to scattering in coupling light from a lithium niobate waveguide to a detector array formed on silicon whether or not channel waveguides are present. Also, accurate alignment of two structures will always be required. These aspects could be eliminated if layers of single crystal silicon could be grown on lithium niobate. Such growth by conventional epitaxy is not possible. However, based on recent experimental results involving laser-assisted growth of single crystal silicon layers on amorphous substrates and formation of MOS devices on these layers, it may be possible to form silicon photodetector arrays on lithium niobate (ref 14).

Acousto-optics offers a potential for unique integrated optical signal processing devices which, in part, is due to the huge difference between acoustic and electromagnetic propagation velocities. Currently under study are several novel thin-film acousto-optic and magneto-optic devices with applications to wideband multi-channel information processing, high-speed deflection and switching of a guided light beam, optical time-division multiplexing/de-multiplexing, and tunable optical filtering (ref 15).

HYBRID OPTO-ELECTRONIC PROCESSING

Many physical problems, especially in simulation and control, require the real-time solutions to 2D and 3D partial differential equations. Under investigation is a hybrid optical-electronic computer capable of high speed solutions of such equations. The speed and versatility of the hybrid computer will enable it to handle a wide variety of complex problems. It will be especially useful in applications requiring partial differential equations to be solved many times while initial conditions, coefficients, nonlinearities, and nonhomogeneous terms are varied. A confocal Fabry-Perot interferometer forms the basis of the optical portion of the computer which performs the Fourier transformations, the spatial filtering, and the feedback (ref 16).

The digital processor operates in an analysis mode with regard to a digitized version of the optical output and acts as a controller for the mirror position and the laser beam scanner which produces the input image and the spatial filters in real time. The feedback configuration allows the optical system to achieve a filtering dynamic range of approximately twenty times that of a more conventional system without feedback. However, even better performance is now being sought through the addition of gain to the system. By offsetting system losses, even modest amounts of gain incorporated into the Fabry-Perot can lead to dramatic increases in filtering dynamic range. Furthermore, sufficient gain would allow the system to operate as an optical operational amplifier. The technique being investigated involves flashlamp-pumped dye lasing. The dye amplifier is placed within the Fabry-Perot and used to offset losses by coherently amplifying a He-Ne signal. In actuality, a two-stage injection-locked dye laser/dye amplifier arrangement is employed to aid in preserving signal frequency and relative phase which is so important in the feedback system.

The Air Force is not interested at this time in supporting investigations into optical logic or digital computing, but rather those areas in which optics can best complement electronic processing. One such area would be where massive parallelism is desirable to perform computationally intensive operations as, for example, in matrix inversion or eigenvector decomposition. Requirements for matrix inversion, such as solutions of simultaneous equations and various signal filtering operations, have established a need for a fast parallel computational method. The problem of real-time eigenvector and eigenvalue determination from a given matrix is also important in practice. Such operations play a role in pattern recognition. Another area which strongly complements electronic processing is the use of optics in VLSI interchip and intrachip communications. Optics may offer a technique for overcoming a fundamental limitation to device density—that of being able to further reduce the spacing between interconnects without appreciably increasing interference. The practical problems of using optics can be separated into two categories: getting suitable optical sources and detectors onto a chip, and realizing suitable holographic optical elements.

REFERENCES

1. C. Warde, A.D. Fisher, J.I. Thackara, and A.M. Weiss, "Image Processing Operations Achievable with the Microchannel Spatial Light Modulator," SPIE Vol. 252- Smart Sensors II, 25-32 (1980).

2. D. Casasent and D. Psaltis, "Multiple-Invariant Space - Variant Optical Processors," Applied Optics 17, 655-659 (1978).

3. R. Kasturi, T. Krile, and J.F. Walkup, "Multiplex Holography for Space-Variant Processing: A Transfer Function Sampling Approach," Applied Optics 20, 881-6 (1981).

4. A. Armand, A.A. Sawchuk, T.C. Strand, D. Boswell, and B.H. Soffer, "Real-Time Parallel Optical Analog-to-Digital Conversion, "Optics Letters 5 (No. 3), 129-31 (1980).

5. B.H. Soffer, D. Boswell, A.M. Lackner, A.R. Tanguay, T.C. Strand, and

A.A. Sawchuk, "Variable Grating Mode Liquid Crystal Device for Optical Processing," SPIE Vol. 218 - Devices and Systems for Optical Signal Processing, 81-7 (1980).

6. B.R. Hunt and S.D. Cabrera, "Optical Computations for a Spatially Adaptive Image Data Compression System," Optical Engineering 20 (No. 4), 616-20 (1981).

7. G. Eichmann, R. Mammone, R. Stirbl, and J. Keybl, "Hybrid Image Source Encoding for Real-Time Image Transmission," SPIE Vol. 154, 204-9 (1978).

8. S.K. Case, "Wavelength Coding for Image Transmission Through Aberrating Media," Optics Letters 6 (No. 7), 311-3 (1981).

9. E.N. Leith and G.J. Swanson, "Achromatic Interferometers for White Light Optical Processing and Holography," Applied Optics 19 (No. 4), 638-44 (1980).

10. G.M. Morris and N. George, "Matched Filtering Using Band-Limited Illumination," Optics Letters 5 (No. 5), 202-4 (1980).

11. G. Gerhart, P.H. Ruterbusch, and F.T.S. Yu, "Color Encoding of Holographic Interferometric Fringe Patterns with White Light Processing," Applied Optics 20 (No. 17), 3085-8 (1981).

12. W.S.C. Chang and P.R. Ashley, "Fresnel Lenses in Optical Waveguides," IEEE Journal of Quantum Electronics QE-16 (No. 7), 744-54 (1980).

13. C.M. Verber, R.P. Kenan, and J.R. Busch, "Correlator Based on an Integrated Optical Spatial Light Modulator," Applied Optics 20 (No. 9), 1626-9 (1981).

14. J.T. Boyd, S. Dutta, H.E. Jackson, and A. Naumaan, "Reduction of the Effects of Scattering by Laser Annealing of Optical Waveguides and by Use of Integrated Waveguide Detection," SPIE Vol. 269 - Integrated Optics, 125-8 (1981).

15. C.S. Tsai, "Guided-Wave Acousto-Optics - Fundamentals and Wideband Applications," SPIE Vol. 139 - Guided Wave Optical Systems and Devices, 132-8 (1978).

16. J. Cederquist and S.H. Lee, "Confocal Feedback Systems with Space Variance, Time Sampling, and Secondary Feedback Loops," Journal of the Optical Society of America 71 (No. 6), 643-50 (1981).

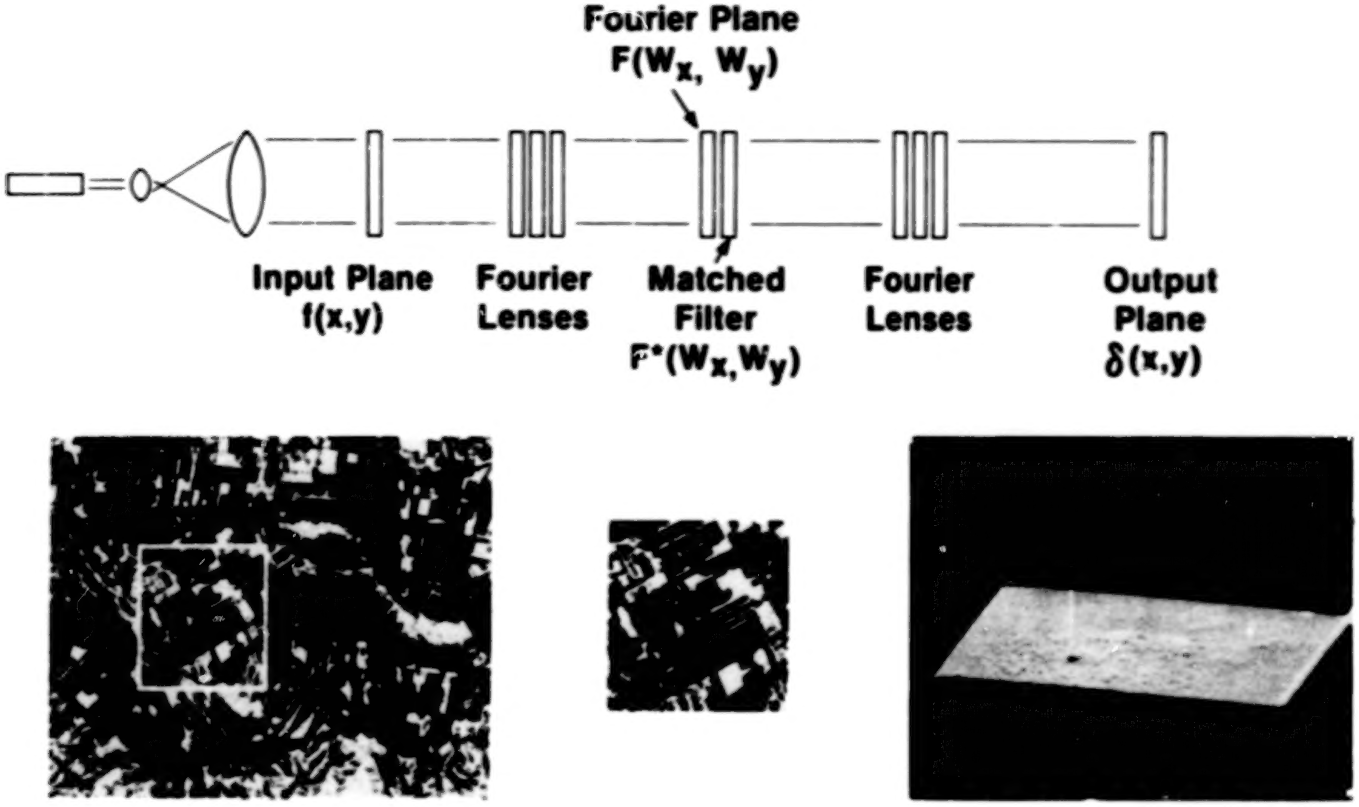
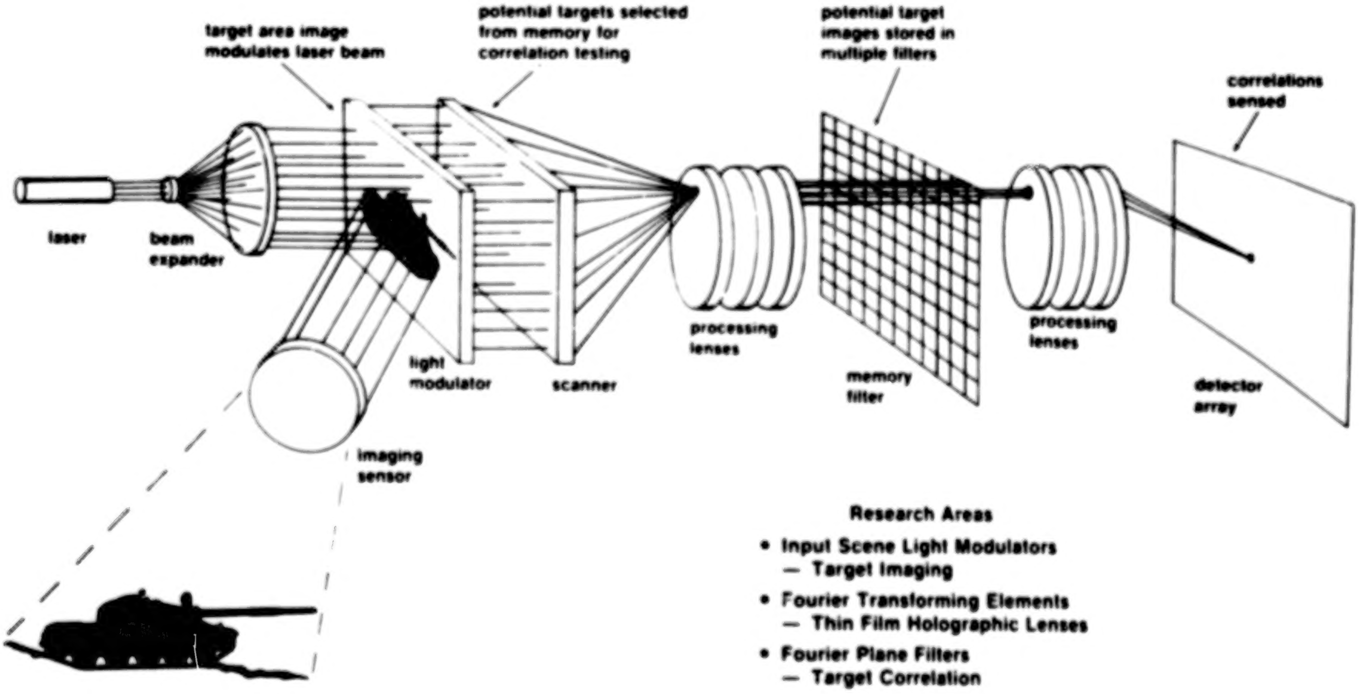

Fourier Plane
F(Wx, Wy)
Input Plane
f(x,y)
Fourier Lenses
Matched Filter
F*(Wx,Wy)
Fourier Lenses
Output Plane
δ(x,y)

Figure 1.- Optical matched filter for real-time scene matching with output shown as a computer-generated plot of light intensity.


target area image modulates laser beam
potential targets selected from memory for correlation testing
potential target images stored in multiple filters
correlations sensed
laser
beam expander
light modulator
scanner
processing lenses
memory filter
processing lenses
detector array
imaging sensor

Research Areas

- Input Scene Light Modulators
  — Target Imaging
- Fourier Transforming Elements
  — Thin Film Holographic Lenses
- Fourier Plane Filters
  — Target Correlation

Figure 2.- Optical matched filter used in pattern recognition mode.

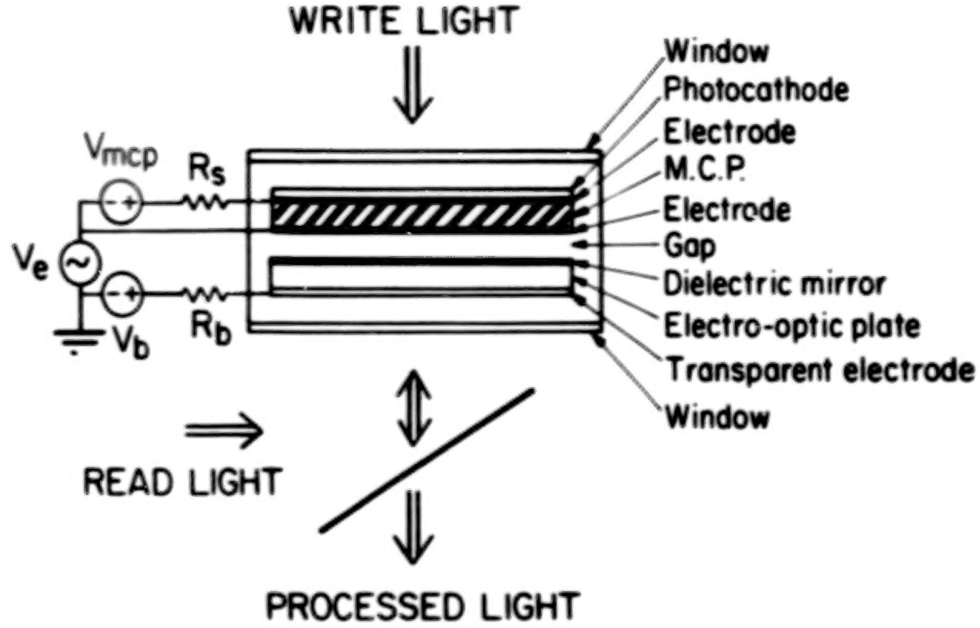

WRITE LIGHT
Vmcp
Rs
Ve
Vb
Rb
Window
Photocathode
Electrode
M.C.P.
Electrode
Gap
Dielectric mirror
Electro-optic plate
Transparent electrode
Window
READ LIGHT
PROCESSED LIGHT

Figure 3.- Microchannel spatial light modulator.

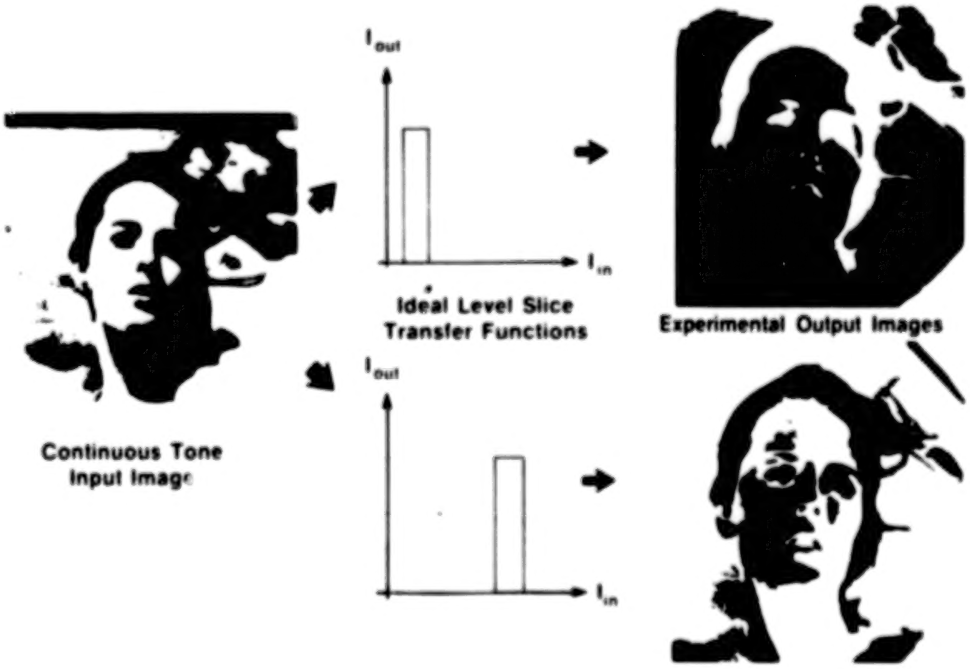

Continuous Tone Input Image
I_out
I_in
Ideal Level Slice Transfer Functions
I_out
I_in
Experimental Output Images

Figure 4.- Level slice operation.

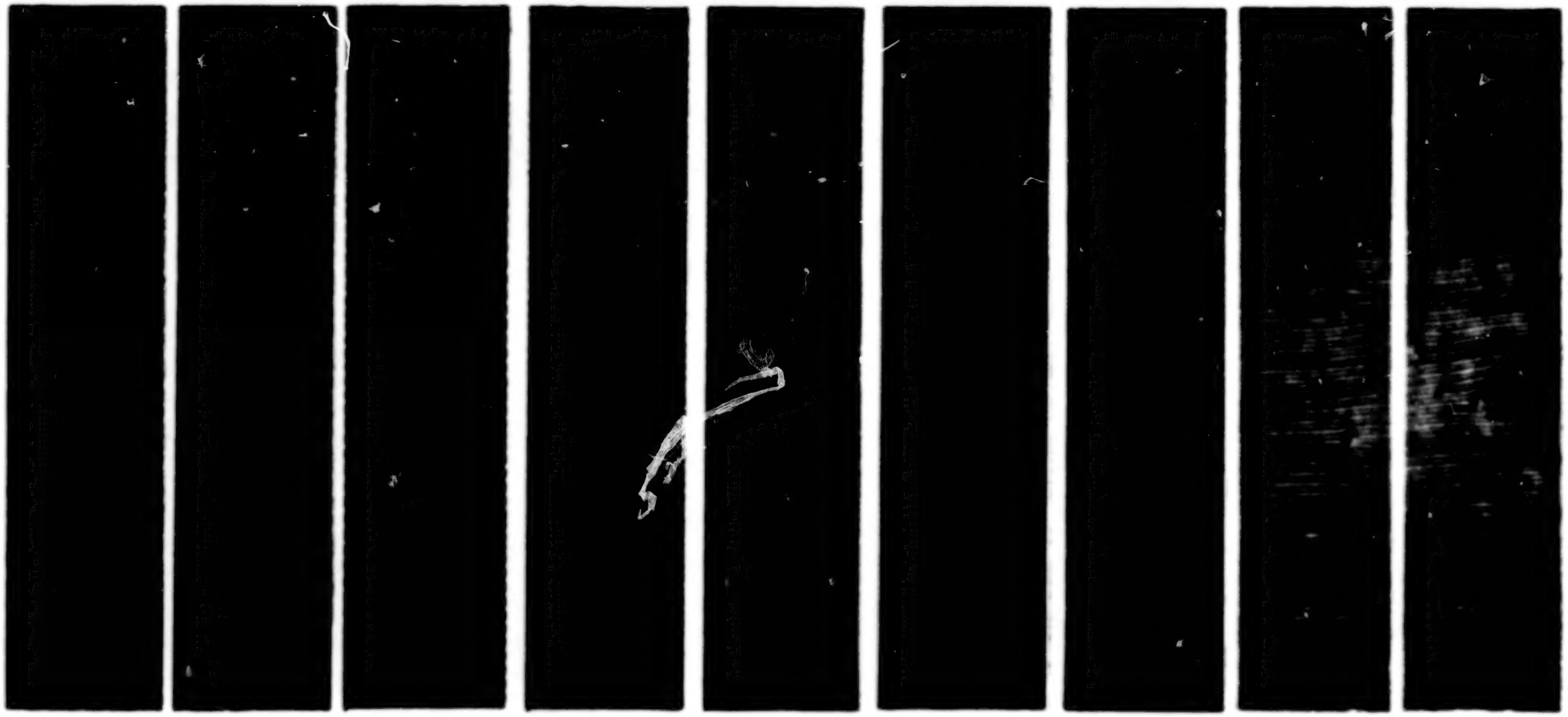

20V
30V
40V
50V
60V
70V
80V
90V
100V

Figure 5.- Variable grating effect viewed through a polarizing microscope.

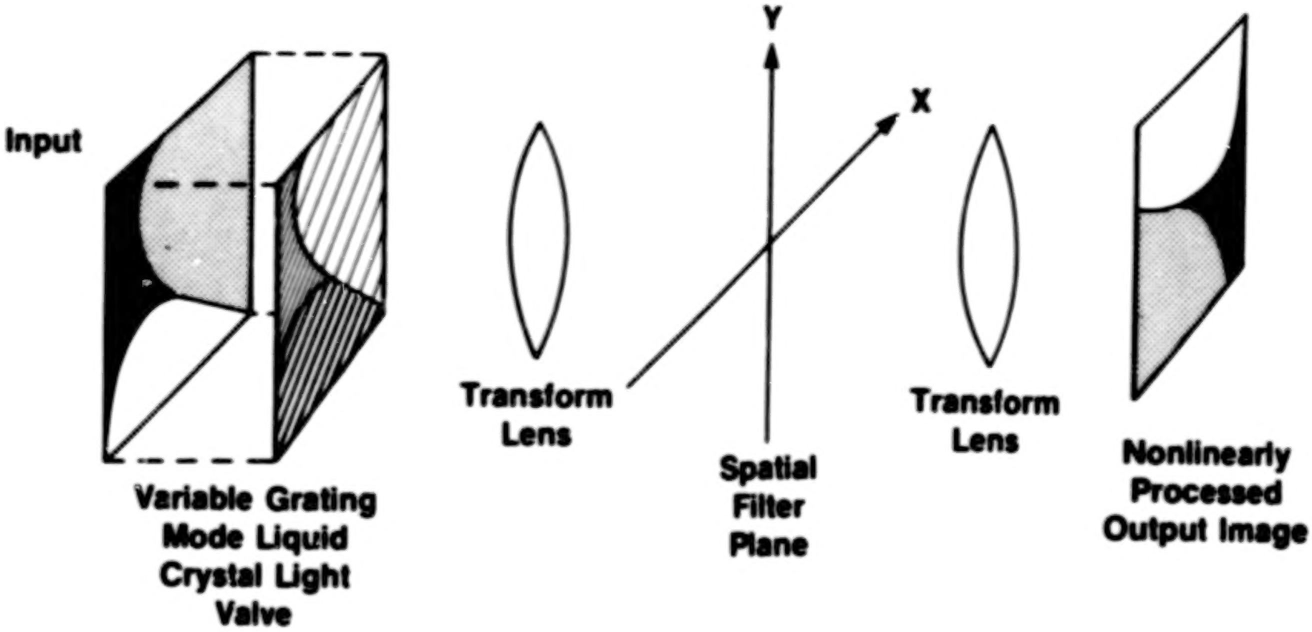

Input
Y
X
Transform Lens
Spatial Filter Plane
Transform Lens
Variable Grating Mode Liquid Crystal Light Valve
Nonlinearly Processed Output Image

Figure 6.- Variable grating mode image processing.

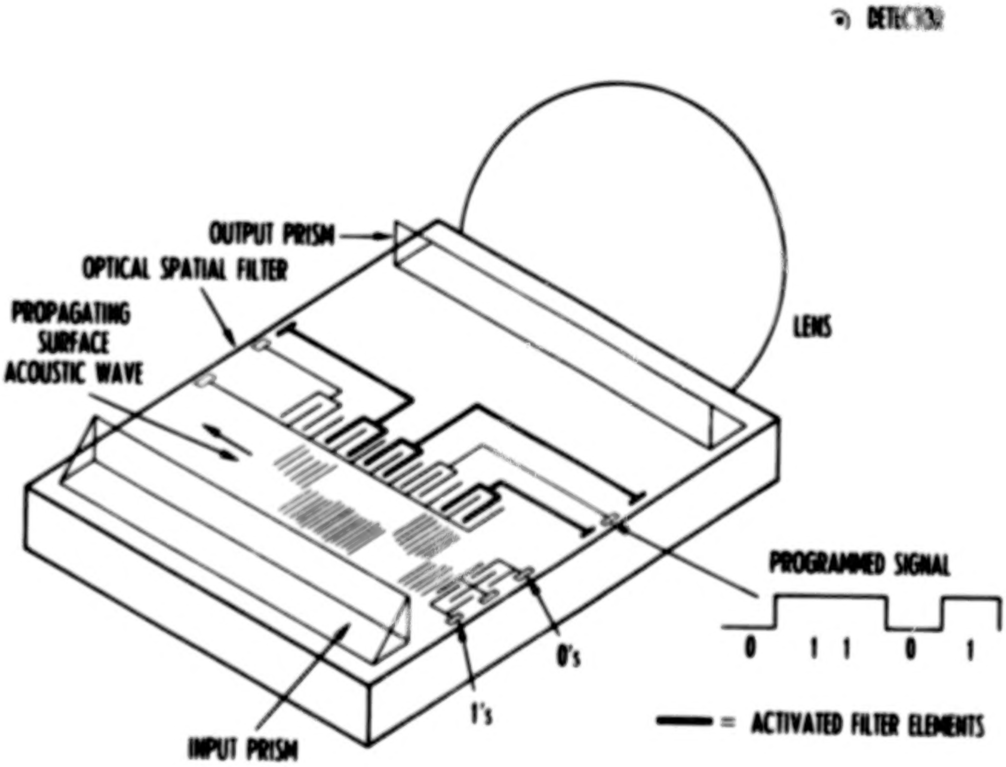

DETECTOR
OUTPUT PRISM
OPTICAL SPATIAL FILTER
PROPAGATING SURFACE ACOUSTIC WAVE
LENS
PROGRAMMED SIGNAL
0 1 1 0 1
0's
1's
= ACTIVATED FILTER ELEMENTS
INPUT PRISM

Figure 7.- Programmable optical waveguide spatial filter.

EARTH FEATURE IDENTIFICATION FOR ONBOARD MULTISPECTRAL DATA EDITING: COMPUTATIONAL EXPERIMENTS

R. M. Aherron
Information & Control Systems, Inc.*

R. F. Arduini
Computer Sciences Corporation*

R. E. Davis, F. O. Huck, S. K. Park
NASA Langley Research Center*

ABSTRACT

A computational model of the processes involved in multispectral remote sensing and data classification is being developed as a tool for designing smart sensors which can process, edit and classify the data that they acquire. An evaluation of sensor system performance and design tradeoffs can be expected to involve classification rates and errors as a function of number and location of spectral channels, radiometric sensitivity and calibration accuracy, target discrimination assignments, and accuracy and frequency of compensation for imaging conditions. This model seeks to provide a link between the radiometric and statistical properties of the signals to be classified and the performance characteristics of electro-optical sensors and data processing devices. Preliminary computational results are presented which illustrate the editing performance of several remote sensing approaches.

INTRODUCTION

To overcome present inefficiencies in worldwide monitoring of resources and the environment by remote multispectral sensing, it is necessary to develop multispectral sensor systems which are "smart" enough so that they can be relied upon to perform such tasks as identifying and locating features of interest, editing out areas of extensive cloud cover and haze, and compensating for atmospheric variability. The development of such smart-sensor systems must take into account the complex natural variability of surface and cloud reflectance and atmospheric radiative transfer. To do so, smart-sensor concepts should be developed and evaluated first as models in the computer, and only thereafter, if promising, as actual devices and systems.

A comprehensive computational model of the deterministic and stochastic processes involved in remote sensing is currently being developed as such a tool for studying multispectral sensor systems and concepts.[1] This model accounts for remote multispectral data acquisition and processing as a function of both deterministic and stochastic elements of solar irradiance, atmospheric radiative transfer, surface and cloud reflectance, and sensor response. The model differs from other related efforts[2-7] in two aspects: One, it treats all elements of the remote sensing process as parts of a single system. Two, it specifically relates stochastic properties of the sensor signal to stochastic properties of the atmospheric radiative transfer and scene spectral reflectance.

---

*Located in Hampton, Virginia.

In this paper we use the computational model of the remote sensing process to study earth feature identification algorithms for onboard data editing. The objective of the feature identification task is to distinguish between vegetation, bare land, water, clouds and snow. If such discrimination between these categories could be made with reasonable accuracy and computational simplicity, then onboard data processing could be relied upon to reduce drastically the amount of data that needs to be transmitted and processed for routine remote sensing operations.

The analysis in this paper is limited to two spectral channels which are centered at wavelengths 0.65 and 0.85 μm. These spectral channels are particularly well-suited for distinguishing vegetation from other earth surface features. For that reason they have been selected for the Feature Identification and Location Experiment (FILE) on Shuttle-OSTA 1. These two channels correspond closely to channels 3 and 4 of the Thematic Mapper (TM) and Multispectral Linear Array (MLA)[8] so that the results presented in this paper are relevant to data processing studies concerned with future remote sensor systems for monitoring earth resources and the environment. Results presented in this paper compare the relative accuracy and computational complexity of three decision techniques for performing the feature identification task. The driving variables are atmospheric conditions, solar incidence angle, and spectral reflectance properties. No attempt is made here to distinguish between clouds and snow since clouds are most efficiently identified, and discriminated from snow and ice, in the reflected IR around 1.5 μm and the emitted IR around 10 to 12 μm (atmospheric window).[9-10]

REMOTE SENSING MODEL

In this section we briefly review the computational remote sensing model developed by Huck et al.[1] The model accounts, as depicted in Fig. 1, for data acquisition and classification. Data acquisition must account for the solar irradiance, atmospheric radiative transfer, surface reflectance, and sensor response. In mathematical terms, data classification is that process that maps the very large sensor sample space into a much smaller space of predefined categories or features. It is essentially the feature identification algorithm that defines these categories.

Radiative Transfer

Deterministic Processes - When atmospheric attenuation (absorption and scattering) and Lambertian ground reflectance effects are taken into account, the radiance derived from solar spectral irradiance incident on a downward-looking sensor is represented as:

$$L = \frac{1}{\pi} E_o T_o \mu_o \rho T + L_d \rho T + L_p \tag{1}$$

where $E_o \equiv E_o(\lambda)$ is the solar spectral irradiance at the top of the atmosphere; $T_o \equiv T_o(\lambda,\tau,\mu_o)$ is the atmospheric transmittance along the incident path from the sun to the surface (solar zenith angle = θ_o, $\mu_o = \cos\theta_o$); $L_d \equiv L_d(E_o,\lambda,\tau,\mu_o,\rho)$ is the diffuse sky spectral radiance which results from all radiation scattered downward onto the surface (i.e., integrated at the target over elevation and azimuth); $\rho \equiv \rho(\lambda)$ is the spectral reflectance of the surface (sometimes called "signature"); $T \equiv T(\lambda,\tau,\mu)$ is the atmosphere transmittance along the exitant path from the surface to the sensor (zenith angle = θ, $\mu = \cos\theta$); and $L_p \equiv L_p(E_o,\lambda,\tau,\mu_o,\mu,\phi)$ is the path spectral radiance which results from all radiation scattered upward along the path from the surface to the sensor. The other parameters are wavelength, λ, optical thickness of the atmosphere $\tau \equiv \tau(\lambda)$, and azimuth angle ϕ between the planes of incidence and exitance. The component of the total radiance L which arises from radiation reflected from the target is referred to as the beam spectral radiance L_b, that is, $L_b = L - L_p$ and $L_b \equiv L_b(E_o, \lambda, \tau, \mu_o, \mu, \rho)$.

The optical thickness $\tau \equiv \tau(\lambda)$, which governs atmospheric transmittance, sky and path radiance, is given by

$$\tau = \sum_{i=1}^{N} \alpha_i \; x_i \tag{2}$$

where $\alpha_i \equiv \alpha_i(\lambda)$ is the attenuation coefficient of the ith atmospheric constituent and x_i the associated attenuator amount (often denoted u_i in the radiative transfer literature). The atmospheric transmittance over the incident path is given by $e^{-\tau/\mu_o}$, and that over the exitant path by $e^{-\tau/\mu}$. The rigorous treatment of multiple scattering which leads to sky and path radiance is very difficult and computationally expensive. To use documented atmospheric radiative transfer models and to keep computations with the deterministic/stochastic model economic. ., we use the AFGL LOWTRAN 4[11] model to account for attenuation in atmospheric absorption bands and the ERIM model (developed and described by Turner[3-4]) to account for scattering. For the spectral region extending from 0.4 to 1.0 µm, the atmospheric radiative transfer is primarily affected by Rayleigh scattering (by air molecules N_2 and O_2), scattering by aerosols (haze droplets and dust), and absorption by ozone (O_3), water vapor (H_2O), and molecular oxygen (O_2).

The ERIM radiative transfer model is used to account for single and multiple scattering in the atmosphere for a surface with a simple geometric pattern. We assume a horizontally homogeneous atmosphere which is bounded by a surface that consists of an infinitesimally small target and a large surrounding background, both of which have uniform diffuse (Lambertian) reflectances. The model accounts for realistic anisotropic phase functions and vertical changes in the atmosphere and for attenuation by ozone absorption and aerosol and molecular scattering but not for attenuation by water vapor and molecular oxygen. This limitation constrains the application of the ERIM model to the evaluation of spectral responses that do not significantly overlap the major H_2O (0.86 to 0.99 µm) and O_2 (0.75 to 0.77 µm) absorption bands. More comprehensive computational models are available but their use as subroutines would be too expensive.

The difference between the LOWTRAN 4 and ERIM models can be expressed in terms of Equation 1. Both models calculate the first term in the same way. LOWTRAN 4 does not, however, include either the second or the third terms which represent scattered solar radiation.

Stochastic Processes - We regard the spectral radiance $L \equiv L(\lambda)$ that reaches the sensor as a stochastic process whose value at each wavelength depends upon a number of random variables associated with both the atmosphere and surface. Letting the operation $E\{\cdot\}$ denote the expectation (average) taken over the ensemble of all possible radiances associated with a particular surface, the mean $\langle L(\lambda)\rangle$ and autocovariance $C_L(\lambda, \lambda')$ of the radiance can be expressed as

$$\langle L(\lambda)\rangle = E\{L(\lambda)\}$$

and

$$C_L(\lambda, \lambda') = E\{[L(\lambda) - \langle L(\lambda)\rangle][L(\lambda') - \langle L(\lambda')\rangle]\}$$

To simulate the effects of atmospheric variability, we assume that the attenuator amounts x_i in Equation 2 are random variables with a known mean x_i and covariance σ_{ij}.

It is particularly convenient, although not necessary for the purpose of simulation, to assume that the vector of attenuator amounts is multivariate Gaussian. Therefore the optical thickness will have a log-normal distribution.

To simulate the effects of surface reflectance variability, we model the reflectance of a particular target surface by

$$\rho(\lambda) = \rho_o(\lambda)\, e^{-x_o \beta_o(\lambda)}$$

where $\rho_o(\lambda)$ and $\beta_o(\lambda)$ are deterministic functions which are characteristic of the surface, and x_o is a standard normal random variable with mean = 0 and variance = 1. For each surface the parameters $\rho_o(\lambda)$ and $\beta_o(\lambda)$ are estimated from empirical reflectance data.

Signal Vector and Reference Pattern

The sensor converts the radiance L into the signal vector s with components s_j which are represented by the simple model

$$s_j = \int_0^\infty L(\lambda)\, S_j(\lambda)\, d\lambda$$

where $S_j(\lambda)$ is the deterministic spectral response of the jth channel. Effects due to electronic noise and errors in radiometric calibration are not treated here, in order to simplify the following formulations.

Because the radiance is stochastic, the signal vector s is a multivariate random variable whose mean r and covariance C have components denoted by

$$r_j = E\{s_j\}$$

and

$$c_{jj} = E\{(s_j - r_j)(s_{j'} - r_j)\}$$

where both j and j' take on the values 1, 2, ..., J, and J is the total number of channels. It follows from the linearity of the signal conversion process that the reference patterns can be computed as

$$r_j = \int_0^\infty \langle L(\lambda)\rangle\, S_j(\lambda)\, d\lambda$$

and

$$c_{jj} = \int_0^\infty \int_0^\infty C_L(\lambda,\lambda')\, S_j(\lambda)\, S_{j'}(\lambda')\, d\lambda\, d\lambda'$$

RADIOMETRIC PROPERTIES AND SENSOR RESPONSES

In this section we present typical radiometric properties encountered in remote sensing and the spectral responses of the FILE sensor system. Our characterizations here are confined to a spectral region of 0.4 to 1.0 µm.

Model Inputs

Solar irradiance - The solar spectral irradiance $E_o(\lambda)$ incident upon the top of the atmosphere is relatively well known and is shown in Figure 2. Its variability is small compared to other uncertainties and thus is ignored here.

Surface reflectance - Table 1 summarizes the categories and substances used as examples in this investigation as well as the (assumed) standard deviations of their reflectances. The spectral reflectance curves are shown in Figure 3. Typical variability of spectral reflectances realized in the simulation is shown in Figure 4.

Atmospheric properties - The mean whole-atmosphere (i.e., vertical path) attenuator amounts for each of these atmospheric constituents, along with reasonable values of their standard deviations based on estimates of climatological variability, are listed in Table 2. All attenuator amounts are assumed to be uncorrelated, except for water vapor and aerosol which are assumed to have a perfect positive correlation. Figure 5 shows a plot of typical average radiance components. Figure 6 shows a typical realization of simulated radiance variability using the LOWTRAN 4 model.

Sensor response - The FILE system consists of two sensors, one centered at .65 µm and one centered at .85 µm. These are relatively narrow-band sensors spectrally (20 nm bandwidth), compared to the 100 to 300 nm bandwidth used in the LANDSAT sensor system. The sensor bandwidths do not overlap any major atmospheric absorption bands. The FILE system includes an onboard categorizer that assigns each pixel to bare land, vegetation, water, or snow/clouds. The FILE sensor channels were originally chosen because of the desirable properties of the ratio of the two channel outputs. Also, channel ratios have some useful compensation effects for sun angle and atmospheric variability.[12]

FEATURE IDENTIFICATION TECHNIQUES

In this section we describe three methods that could be used to identify (edit) such categories as vegetation, bare land, water, and cloud/snow. Two of these methods are closely related; they use either the maximum likelihood (MLH) or mean-square distance (MSD) classification algorithms to classify the signal vectors according to reference patterns (or training sets) and then collect all classes into categories. The third method, referred to as the boundary approximation method (BAM), avoids the classification step and assigns the signal vectors directly into categories.

Classification

A very common classification procedure which makes use of covariance information is based upon maximizing the (assumed) Gaussian PDF. Specifically, a particular signal vector, s, representing J spectral channels, is classified as spectral type n, provided the Gaussian PDF of s conditioned on spectral type n is largest, i.e., provided

$$\mathrm{PDF}(s|n) = \frac{1}{(2\pi)^{J/2}|C_n|^{1/2}} \exp\left[-\frac{1}{2}(s-r_n)^t C_n^{-1} (s-r_n)\right] \qquad (3)$$

is greater than PDF $(s|n')$ for all other spectral types $n' = 1, 2, \ldots N$. To avoid the computational expense of evaluating the exponential, an equivalent classification procedure can be used which minimizes $-\log_e [\mathrm{PDF}(s|n')]$. This, in turn, is equivalent to determining the n which minimizes

$$(s-r_n)^t C_n^{-1} (s-r_n) + \log_e |C_n| \tag{4}$$

where $|\cdot|$ denotes determinant. If the a priori probability of each spectral type is known (and all types are not equally likely), it is often desirable to weight the classification procedure by this prior information. However, no weighting is used for the results presented in this paper. For this classifier, the mean and covariance matrix uniquely determine a class. This is the maximum likelihood classifier (MLH).

Several simplifying assumptions concerning the nature of the covariance matrix, given imperfect knowledge, can be made. In the special case of

$$C_n = I$$

where I is the identity matrix, minimizing Equation 4 becomes equivalant to minimizing

$$(s-r_n)^t (s-r_n)$$

This is commonly called the mean-square distance (MSD) or Euclidean-distance classifier. From a computational standpoint, we eliminate the evaluation of one vector-matrix (lxn * nxn) multiplication and the addition of the logarithm term. For this special case the mean uniquely determines the class.

Editing

Aggregation method - The MLH and MSD classification algorithms work with classes but can be used to distinguish between categories such as vegetation, bare land, water, and clouds/snow by collecting a set of classes to define each of the four categories. For example, if one can pick some small number of vegetation classes that approximate (span) the vegetation category then one has a method of categorizing vegetation. In other words one can blanket the area, in spectral space, occupied by the vegetation category with a set of classes. Computationally, what can be done is to actually assign a signal to a class (according to the chosen decision rule) but to only retain the category within which that class falls. For example, a signal from a corn target might be classified as wheat but the editor retains only the fact that it has been categorized as vegetation. The overall accuracy of categorizing will be higher than the overall classification accuracy because intra-category class confusion is not counted as wrong. We call this method categorization by aggregation.

Boundary approximation method - In discussing this editing technique frequent reference will be made to covariance ellipse plots. These ellipses are defined by the following equation:

$$(x-r_n)^t C_n^{-1} (x-r_n) = 1$$

where x is the channel vector. If the nth class was truly multivariate-Gaussian then this ellipse would enclose 66% of the data. The ellipses are centered about the class mean. These plots provide a very useful approximation to the scatter of the data due to such things as the variability in radiance shown in Figure 6. Figure 7 shows ellipse plots for simulated data for a total of twenty classes. This data represents a simulation for 23 km visual range and a solar zenith angle of 30°, using the FILE channels.

As stated previously this particular set of FILE spectral bands was chosen because of the properties that the channel ratio possessed. Figure 7 shows 3 lines superimposed. The lines are chosen so as to delineate the four categories. These lines represent simple approximations to the boundaries of the categories in spectral space. Hence the name Boundary Approximation Method (BAM). The lines drawn are not strictly defined by the channel ratios. It was decided that, by adding an offset to the two diagonal equations, much better performance could be achieved. This causes only a slight increase in computational costs. The third line represents an "absolute" radiance threshold which forms the boundary between the bare land and snow/cloud, categories which have similar channel ratios. However, this threshold has a sensitivity to solar zenith angle.[12] The three boundaries are defined by the following equations. If x_i is the value of channel i then the three lines are defined as:

$$x_2 = 1.45\ x_1 - 4.1$$
$$x_2 = 3.0\ x_1 - 2.1$$
$$x_2 = 16.5$$

As one can see this method is very simple in terms of the calculations necessary to categorize a signal and is similiar to the standard parallelepiped classifiers. This method represents the opposite end of the spectrum, in terms of computational complexity, from the aggregation method.

Space qualified data processing hardware is still rather expensive and bulky. Any implementation of onboard processing of multispectral data will encounter very conservative limits in terms of available processing and storage capacity. This, in turn, limits the computational "costs" that can be incurred in identifying features. Estimates of the computational complexity of the two methods of aggregation were made after the formulas were simplified and modified as much as possible. No provision was made for special processing architecture (i.e., parallelism) in the estimates. For the boundary approximation technique we have two equations to evaluate with 1 add and 1 multiply each for a total of 4 add/multiplies. The number of add/multiplies per pixel necessary to do the MSD classification, and thus editing, can be approximated by 2LJ where L is the number of classes and J is the number of channels. Similarly, the number for PDF can be approximated by $(L/2)(J^2+3J)$. Those interested in more details might start with Ref. 13 and 14.

COMPUTATIONAL EXPERIMENTS AND RESULTS

Experiments

In this section we present results of computational experiments which illustrate the performance of the MLH and MSD aggregation and the boundary approximation methods in distinguishing between such spectral features as vegetation, bare land, water, and cloud/snow from data obtained with the two FILE spectral channels. These two channels are centered at 0.65 and 0.85 µm and thus approximately coincide with TM/MLA sensor channels 3 and 4. The results provide insight into the susceptibility of feature identification accuracy on the natural variability of atmospheric radiative transfer and target and background spectral reflectance.

The computer program (previously described) simulating the performance of an orbiting multispectral sensor generates pseudo-random observations (i.e., signal vectors) which are assembled into training sets for computing a reference pattern library. Subsequently, observations are generated for input to the editor/classifier which utilizes the reference pattern library or other criteria to make its decision.

Editing accuracy (i.e., the ratio of correct categorizations to the total number of observations) is used as a figure of merit to determine sensor performance. Training sets and editing accuracies are based upon a total of 100 observations per target.

In practice, classification using the elements of a reference library will always be compromised since these elements must be based upon values for the means and covariances obtained at one specific set of conditions. For example, the reference library may have been compiled when the atmosphere was relatively clear, while the observations were made when the atmosphere was hazier. This error source has been included by computing the reference vector r and covariance matrix C for each target at a visual range of about 23 km (representing a moderately clear atmosphere), and the signal vectors s either at the same visual range or at reduced visual ranges of about 10 km and 5 km. Similarily, to assess the effect of changes in solar incidence angle, reference data were obtained for a solar incidence angle of 30°, whereas feature identification decisions were made either at the same or at 40° solar incidence angle. The sensor viewing geometry remained vertical.

Results

Figure 8 summarizes the predicted feature identification accuracy, and Table 3 presents the amount of computations required for the three decision processes. The feature identification accuracies are complex functions of changes in task assignment, visual range, and solar incidence angle. Nevertheless, it can be concluded that the MLH and MSD aggregation methods provide consistently higher feature identification accuracies than the BAM and also tend to be less susceptible to changes in both visual range and solar incidence angle. However, this improvement is gained at a substantial increase in computational and storage requirements. These requirements would increase rapidly for the MLH and MSD aggregation methods with increases in the number of substances to be accounted for and in the number of spectral channels to be used. One can also see that for a 3-fold increase in computational cost little or nothing is gained in terms of increased accuracy or sensitivity to sun angle by using the MLH as opposed to the MSD. The results in figure 8 serve to show that small changes in the mix of classes, observed by the sensor, can have as important an effect as the other sources of error, such as atmospheric degradation. Specifically, ripening barely lacks the strong chlorophyll absorption band that the .65 μm channel is designed to detect in vegetation. Therefore, barley is likely to be confused with bare land.

These results suggest that BAM could become a useful algorithm for spectral feature identification if it is expanded to use more than 2 spectral channels and if changes in solar incidence angle could be accounted for. It would, of course, also be desirable to account for changes in visual range; however, this would be more difficult.

In practice the exact formulation of an algorithm depends on the particular task. For example, in editing out clouds, errors of omission, whereby we retain some cloud data, are less troublesome than errors of commission, whereby we edit out data from other categories. In general, if one particular category is of interest the boundaries for that category can be "relaxed" somewhat in order to pick up marginal outliers. In more complete analysis the a priori probabilities of each of the categories are needed in order to optimize the omission-commission trade-offs. This is also true for the aggregation method.

CONCLUDING REMARKS

Two fundamentally different approaches to feature identification, aggregation and boundary approximation for onboard data editing have been examined. Relatively small differences in accuracies between the simplest boundary approximation and the comparatively complex aggregation method were found, whereas the difference in computational requirements is very large. Further investigation should be performed to determine improvements in feature identification accuracy that may be achieved by careful selection of both the number and location of spectral channels.

The results obtained using the computational model were generally consistent with practical experience with Landsat data in terms of sensitivity to changes in haze, sun angle, and to the set of targets defined. The use of this model as a tool in the preliminary design and evaluation phase of remote sensing systems should prove valuable. To improve computational accuracy it will be necessary to account for the probability of occurence of various targets and the probability of encountering specific atmospheric conditions.

REFERENCES

1. Huck, F. O.; Davis, R. E.; Park, S. K.; Aherron, R. M.; Arduini, R. F.: Computatational Modeling for the Study of Multispectral Sensor Systems and Concepts, Optical Engineering, in press, 1981.
2. Malila, W. A.; Crane, R. B.; Omarzu, C. A.; and Turner, R. E.: Studies of Spectral Discrimination. NASA CR-134181, WRL 31650-22-T, May 1971.
3. Turner, R. E.: Atmospheric Effects in Remote Sensing. Remote Sensing of Earth Resources. ed. F. Shahrokhi, 26-28 March 1973.
4. Turner, R. E.: Atmospheric Effects in Multispectral Remote Sensor Data. NASA CR-ERIM 109600-15-F, May 1975.
5. Kondratyev, K. Ya; Beliavsky, A. I.; and Pokrovsky, O. M.: Possibilities of Optimal Planning of Multipurpose Survey from Space. Proc. of the Thirteenth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, 32-27 April 1979.
6. Landgrebe, D. A.; Biehl, L. L.; and Simmons, W. R.: An Empirical Study of Scanner System Parameters. IEEE Transactions on Geoscience Electronics, Vol. GE-15, No. 3, July 1977.
7. Wiersma, D. J.;and Landgrebe, D. A.: An Analytical Approach to the Design of Spectral Measurements in the Design of Multispectral Sensor. IEEE Machine Processing of Remoetly Sensed Data Symposium, 1979.
8. Salomonson, V. V.: Landsat D, A Systems Overview. Proc. of the Twelfth Inter-National Symposium on Remote Sensing of the Environment, Vol. II, Environmental Research Institute of Michigan, 20-26 April 1978.
9. Valovcin, F. R.: Spectral Radiance of Snow and Clouds in the Near Infrared Spectral Region. AFGL-TR-78-0289, Air Force Surveys in Geophysics, No. 403, November 17, 1978.
10. Stebbins, W. J.: Snow/Cloud Discrimination Staring Mode Radiometer. Electro-Optics/Laser Conference and Exposition, Boston, Sept. 19-21, 1978.
11. Selby, J. E. A.; Kneizys, F. X.; Chetwynd, J. H., Jr.; and McClatchey, R. A.: Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 4. AFGL-TR-78-0053, Environmental Research Papers, No. 626, Feb. 28, 1978.
12. Maxwell, E. L.: Multivariate System Analysis of Multispectral Imagery. Photogrammetric Engineering and Remote Sensing, Vol. 42, No. 9, Sept. 1976.
13. Duda, R. O.; and Hart, P. E.: Pattern Classification and Scene Analysis. John Wiley and Sons, New York, 1973.
14. Benz, H. F.; Kelly, W. L; Husson, C.; Culotta, P. W.; and Synder, W. E.: CCD Data Processor for Maximum Likelihood Feature Classification, Proc. of the AIAA Sensor Systems for the 80's Conference, Colorado Springs, Dec. 2-4, 1980.

Table 1. Targets and the Assumed Standard Deviation of Their Reflectance

Category	Substance	Standard deviation of reflectance, (
I. Vegetation	Barley	.1
	Wheat	.1
	Oats	.1
	Corn	.1
	Aspen	.1
	Red pine	.1
	White pine	.1
II. Bare land	Chernozem-type soil, Nebraska	.125
	Pedialier-type silt, Arkansas	.105
	Pedocal-type soil, Ohio	.105
	Pedocal-type soil, Nebraska	.02
	Quartz sand, Oregon	.140
	Clay, Missouri	.055
	Red quartz and calcite sand, Utah	.105
	Pedocal-type soil, Oklahoma	.09
	Concrete road	.1
	Asphalt road	.1
III. Water	Sea water	.06
IV. Cloud/Snow	Optically thick cloud	.1
	Sugar consistency snow	.08

Table 2. Attenuator Amounts in Vertical Column of Atmosphere

Attenuator	Average value, $\bar{x}$			Standard deviation, σ		
	LOWTRAN	atm·cm	# m^{-2}	LOWTRAN	atm·cm	# m^{-2}
Air (N_2, O_2)	8.0 km	8.0×10^5	2.2×10^{29}	0.24 km	2.4×10^4	6.5×10^{27}
Aerosol: 23 km visual range	1.5 km	1.5×10^5	4.0×10^{28}	0.5 km	0.5×10^4	1.4×10^{28}
10 km visual range	3.0 km	3.0×10^5	8.1×10^{28}	1.0 km	1.0×10^5	2.7×10^{28}
5 km visual range	5.0 km	5.0×10^5	1.4×10^{29}	1.7 km	1.7×10^5	4.6×10^{28}
Ozone (O_3)	0.34 atm·cm	3.4×10^{-1}	9.2×10^{22}	0.10 atm·cm	1.0×10^{-1}	2.7×10^{22}
Water vapor (H_2O)	1.14 g/cm^2	1.14	3.1×10^{23}	0.36 g/cm^2	3.6×10^{-1}	9.7×10^{22}
Molecular oxygen (O_2)	1.7 km	1.7×10^5	4.7×10^{28}	0.12 km	1.2×10^4	3.2×10^{27}

Table 3. Computational Requirements for 2 Spectral Channels

Decision Process	Add/Multiplies with number of targets:	
	17	20
MLH	187	220
MSD	68	80
BAM	4	4

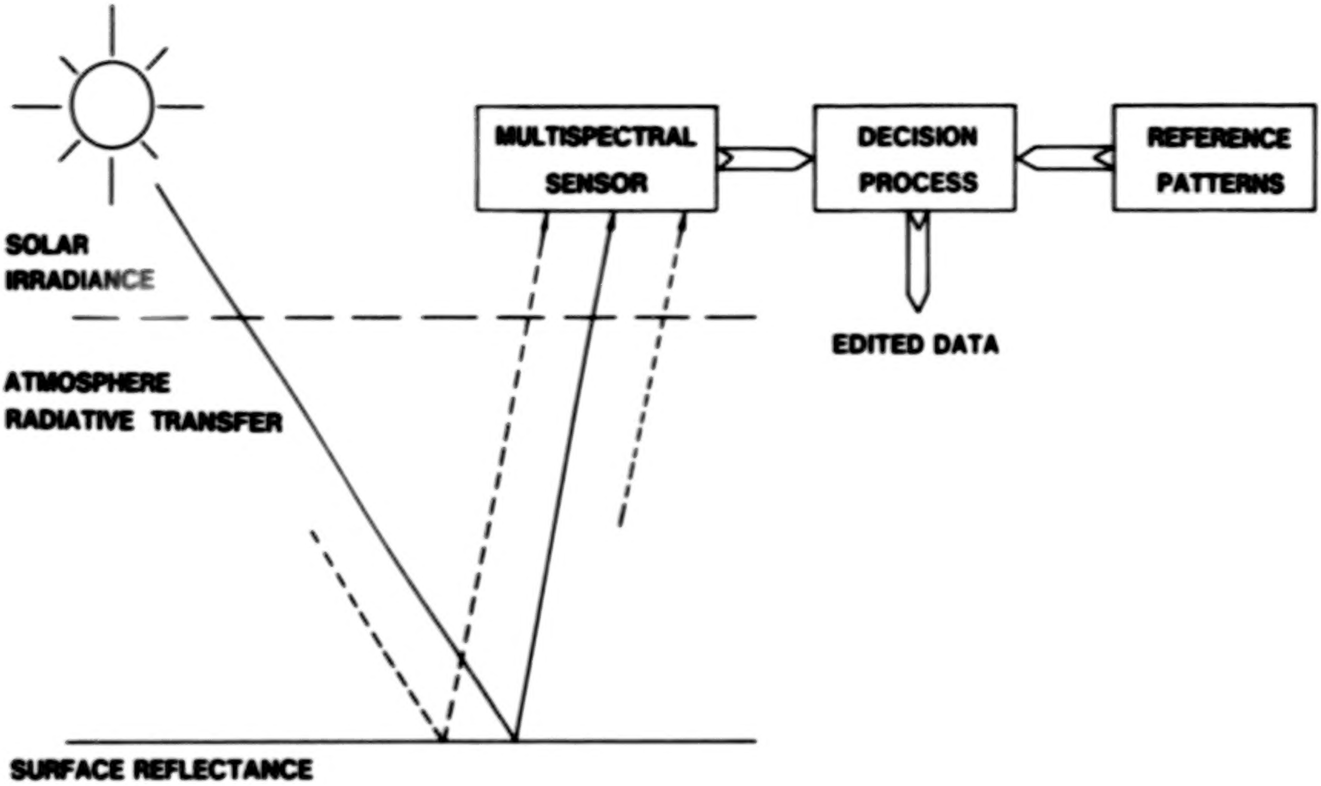

MULTISPECTRAL SENSOR
DECISION PROCESS
REFERENCE PATTERNS
EDITED DATA
SOLAR IRRADIANCE
ATMOSPHERE RADIATIVE TRANSFER
SURFACE REFLECTANCE

Figure 1.- Data acquisition and decision process.

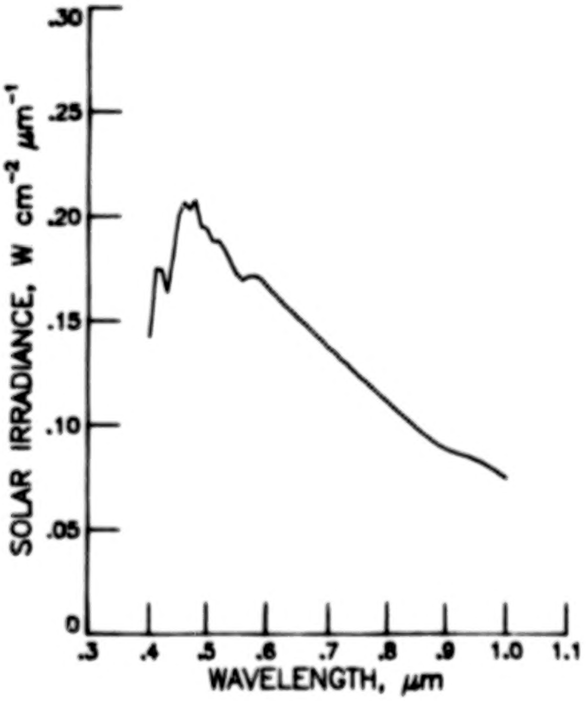


Figure 2.- Solar irradiance at the top of atmosphere.

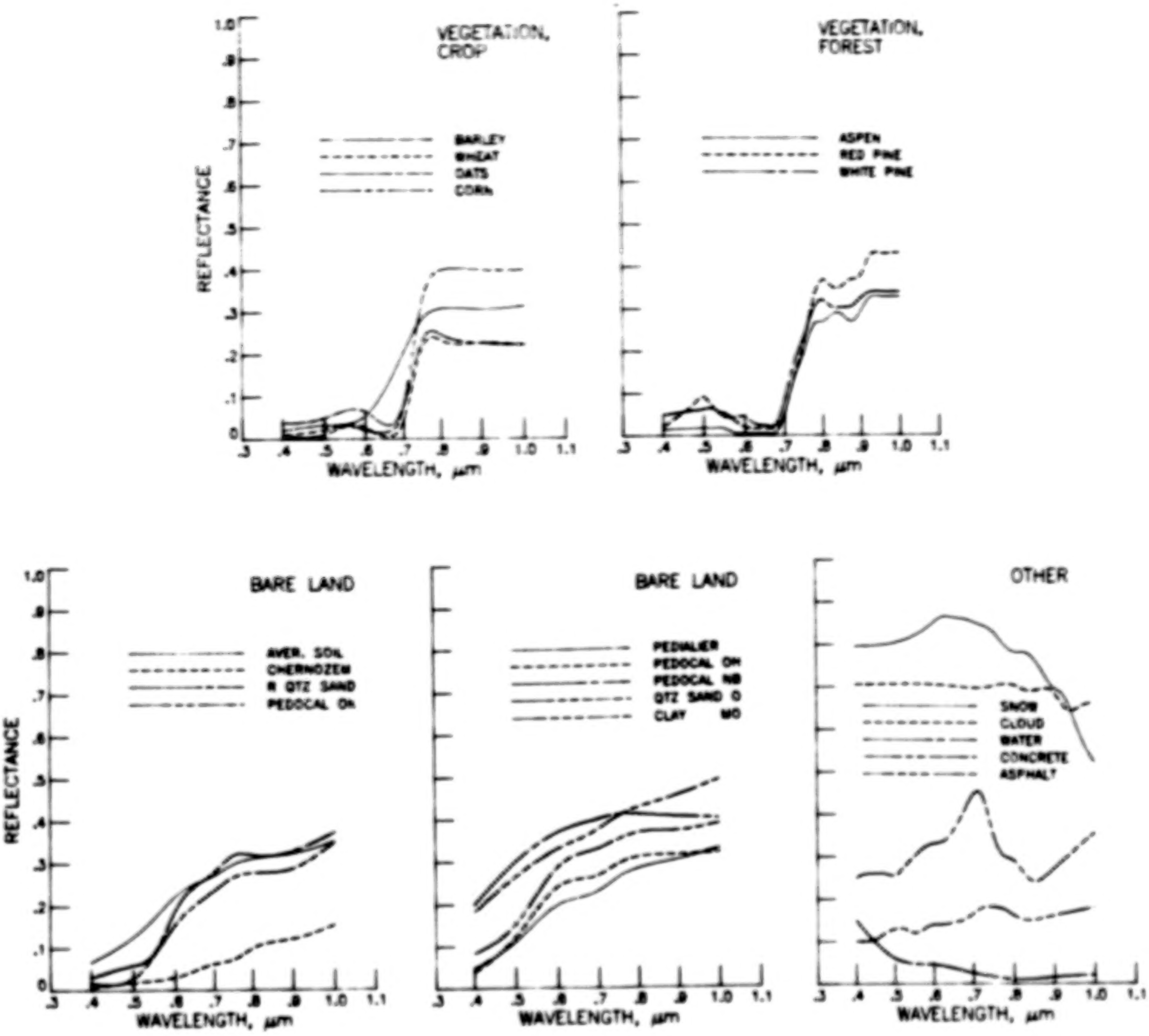


Figure 3.- Spectral reflectances. (See Table 1 for standard deviation of reflectances.)

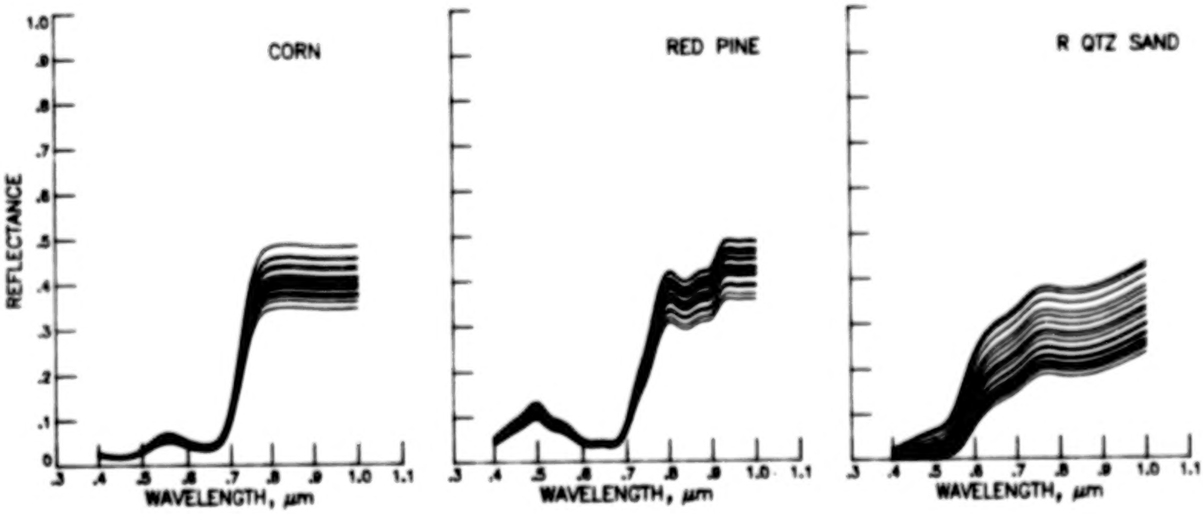


Figure 4.- Typical realization of simulated variability of three spectral reflectances.

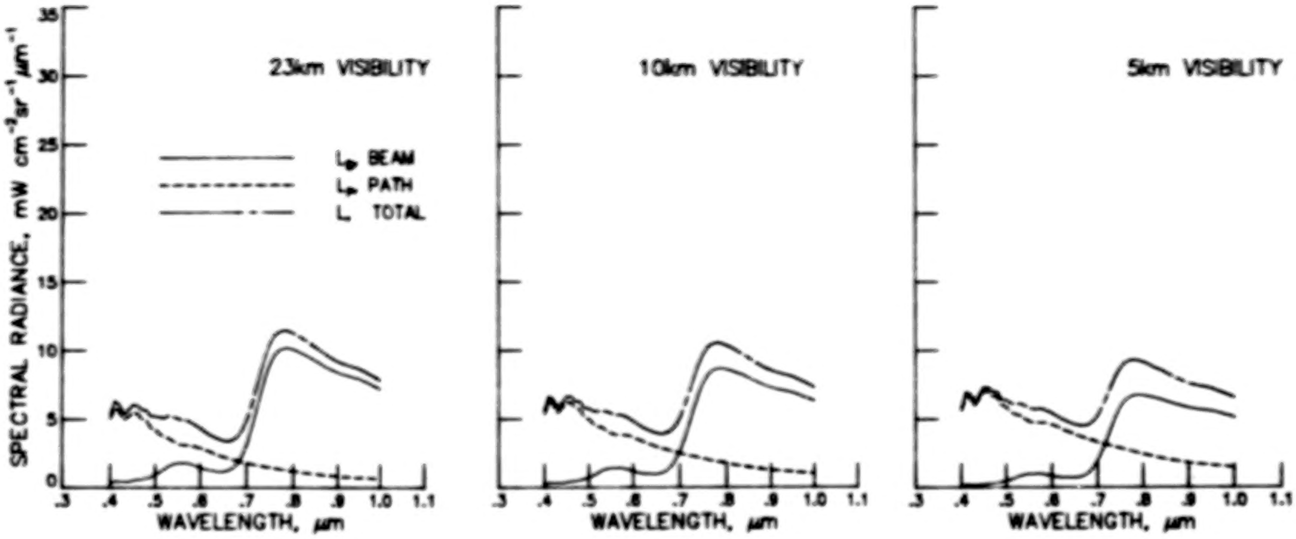


Figure 5.- Average beam, path, and total radiance for corn as target and average soil as background using ERIM radiative transfer model. Exitance is vertical and solar incident angle is 30°.

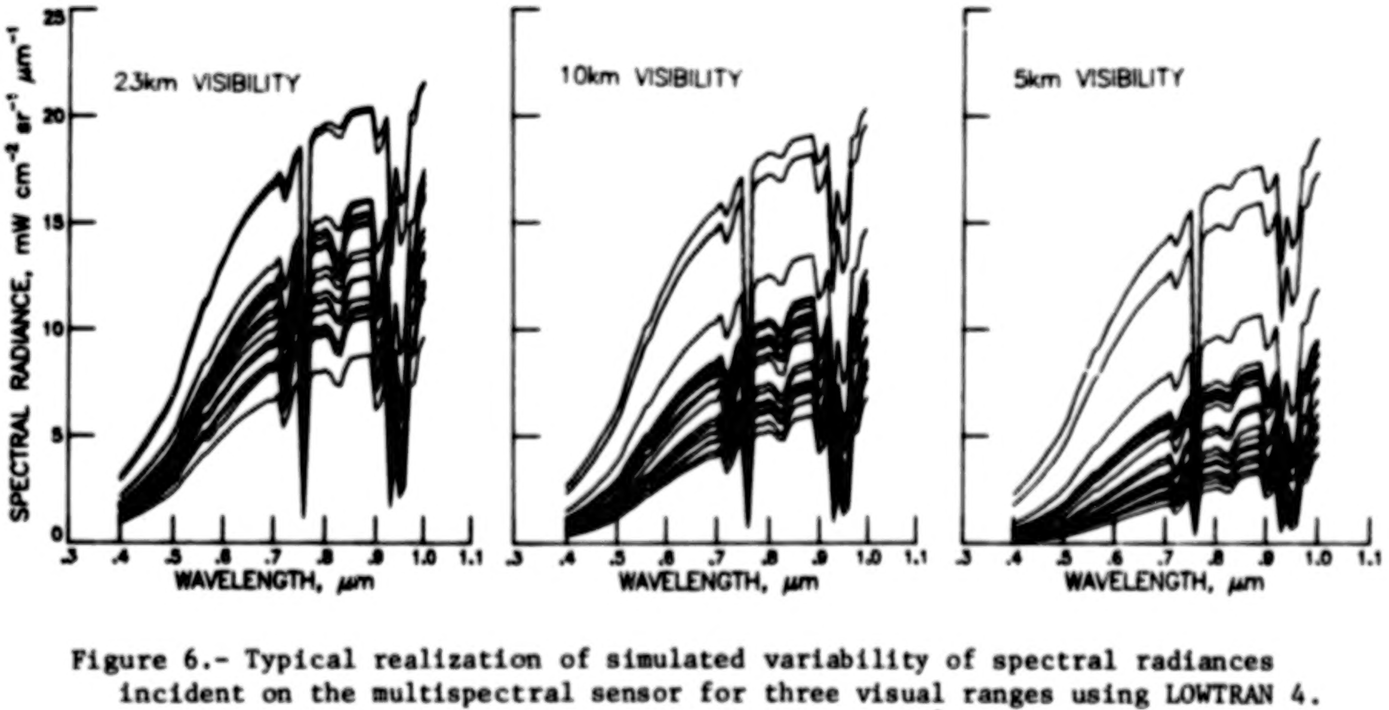
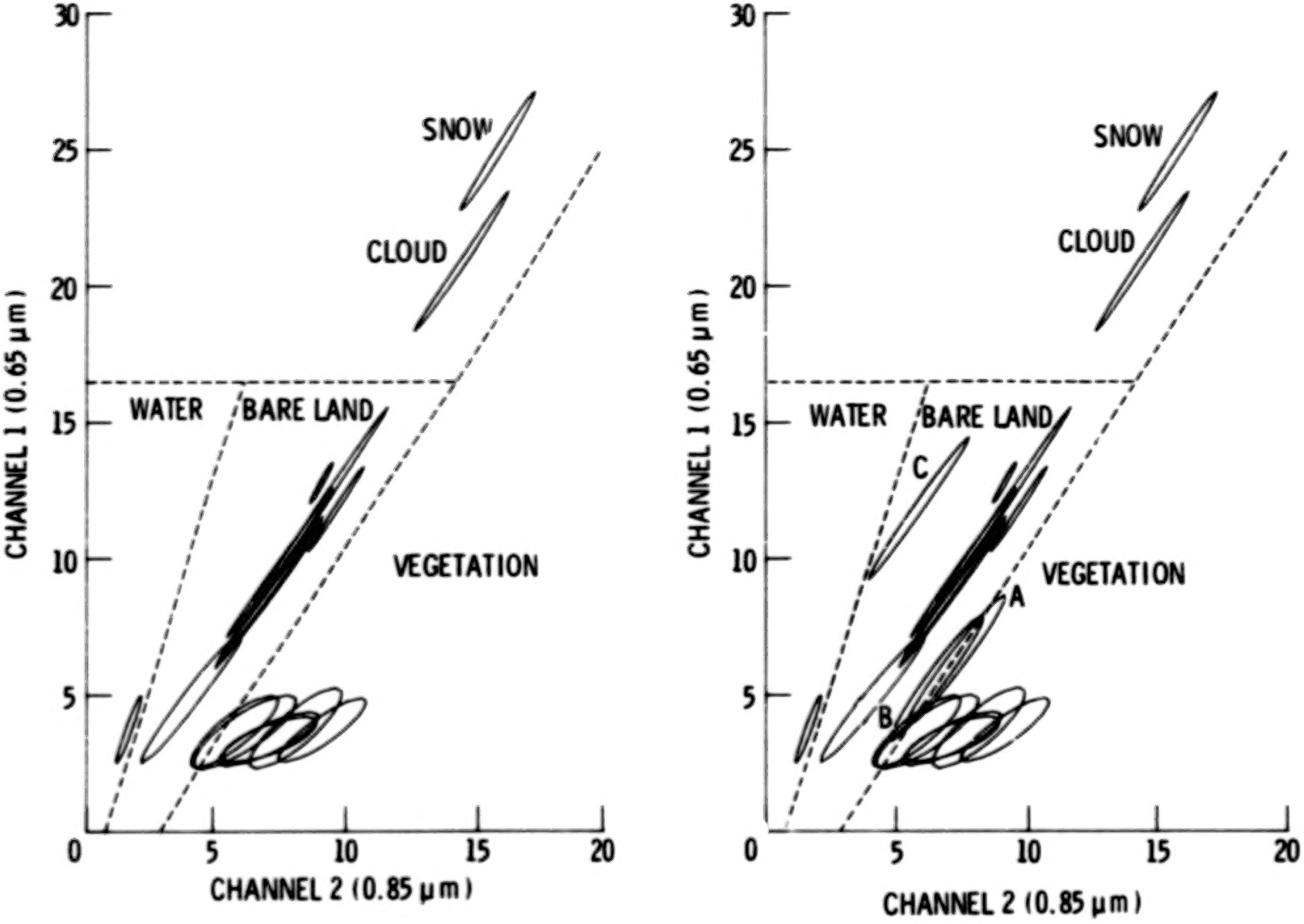


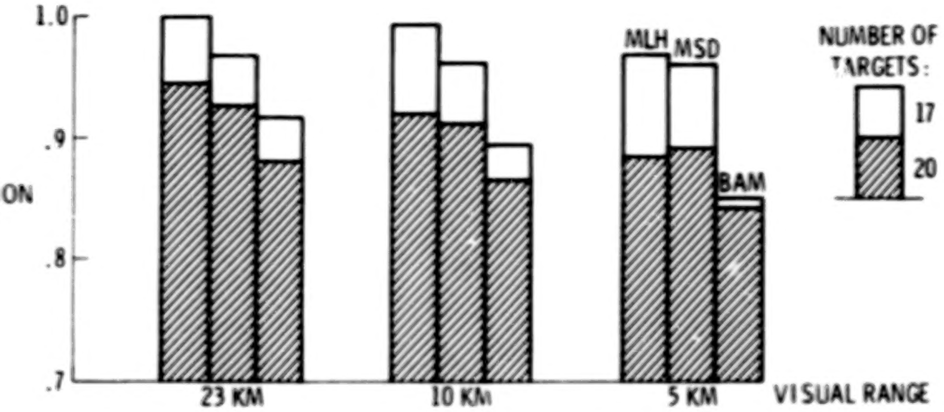
Figure 6.- Typical realization of simulated variability of spectral radiances incident on the multispectral sensor for three visual ranges using LOWTRAN 4. Target is quartz sand and solar incident angle is 30°.



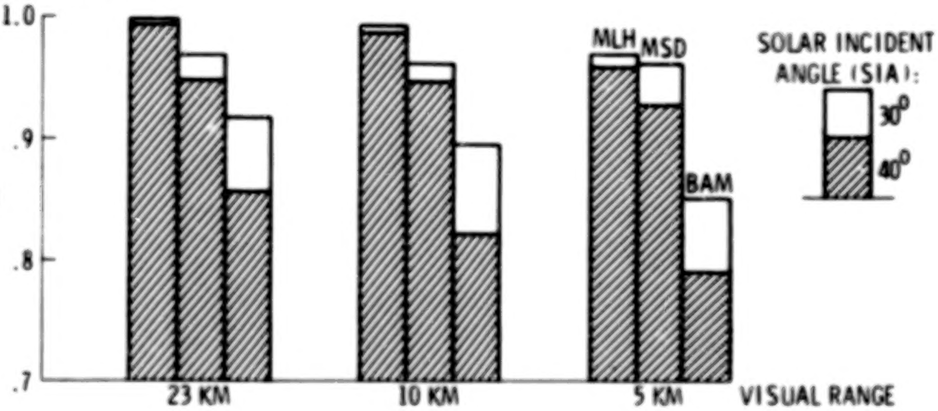
(a) Number of targets is 17; barley, asphalt and concrete are excluded.

(b) Number of targets is 20; barley (B), asphalt (A) and concrete (C) are included.

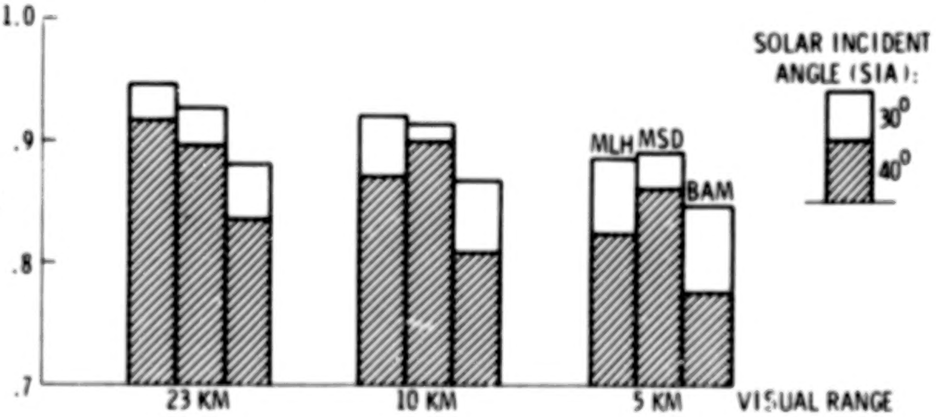
Figure 7.- Covariance plot of signal from two spectral channels for 23 km visual range and 30° solar incidence angle, and regions used for the boundary approximation method (BAM) decision process.



(a) Number of targets is 17 or 20, and data is acquired at a solar incidence angle (SIA) of 30°.



(b) Number of targets is 17, and data is acquired at a SIA of 30° or 40°.



(c) Number of targets is 20, and data is acquired at a SIA of 30° or 40°.

Figure 8.- Feature identification accuracies attained with three decision processes: maximum likelihood (MLH) and mean-square distance (MSD) aggregation method and boundary approximation method (BAM).

ROLE OF OPTICAL COMPUTERS IN AERONAUTICAL CONTROL APPLICATIONS

Robert J. Baumbick
NASA Lewis Research Center
Cleveland, OH 44135

ABSTRACT

More complex aircraft and mission requirements will require larger, faster digital computers to perform a variety of functions. For adaptive control, controller gains are computed as a function of flight conditions. For multi-state problems, high speed computing is required to perform matrix/vector and matrix/matrix operations in order to compute these controller gains.

Continued safe operation of the plant has to be maintained in the event of sensor failure. Accommodation of sensor failures requires a system math model that operates in real-time. The model, used to calculate an estimate of the failed sensed variable, should ideally be a full nonlinear model. Currently simplified models are used because of computing limitations. Because of the above requirements, research is being done to determine the role that optical computers might play in aircraft control. The optical computer has the potential high speed capability required, especially for matrix/matrix operations. The optical computer also has the potential for handling nonlinear simulations in real-time.

In addition to the potential high speed capability of optical computers, they may also be more compatible with fiber optic signal transmission. Fiber optics offer advantages over conventional wire systems. One attractive feature of fiber-optic systems is immunity to noise generated from other electromagnetic sources and immunity to lightning strikes. Optics also permit the use of passive sensors to measure process variables. No electrical energy need be supplied to the sensor. Complex interfacing between optical sensors and the optical computer may be avoided if the optical sensor outputs can be directly processed by the optical computer.

Summary

More complex aircraft and mission requirements will require larger, faster digital computers to perform a variety of functions. For adaptive control, controller gains are computed as a function of flight conditions. For multi state problems high speed computing is required to perform matrix/vector and matrix/matrix operations in order to compute these controller gains.

Continued safe operation of the plant has to be maintained in the event of sensor failure. Accommodation of sensor failures requires a system math model that operates in real-time. The model, used to calculate an estimate of the failed sensed variable, should ideally be a full nonlinear model. Currently, simplified models are used because of computing limitations. Because of the above requirements, research is being done to determine the role that optical computers might play in aircraft control. The optical computer has the potential high speed capability required, especially for matrix/matrix operations. The optical computer also has the potential for handling nonlinear simulations in real-time.

In addition to the potential high speed capability of optical computers, they may also be more compatible with fiber optic signal transmission. Fiber optics offers advantages over conventional wire systems. One attractive feature of fiber-optic systems is immunity to noise generated from other electromagnetic sources and immunity to lightning strikes. Optics also permit the use of passive sensors to measure process variables. No electrical energy need be supplied to the sensor. Complex interfacing between optical sensors and the optical computer may be avoided if the optical sensor outputs can be directly processed by the optical computer.

NASA Lewis Research Center is supporting two grants for basic research on optical computing and how it may be used in aeronautical applications. Carnegie-Mellon University is developing a hybrid electro-optic computer capable of performing matrix operations such as inversion, multiplication, etc. This work considers an analog optical computer.

Ohio State University is developing a digital optical processor. This work centers around a liquid crystal light valve which is used to perform various logic operations and arithmetic operations. Both binary and residue arithmetic are being considered for these applications. Residue arithmetic is potentially more attractive because of the inherent parallelism of this type of arithmetic.

The hybrid and digital optical processors both offer potential improvements in computing speed compared to all-electronic processors, but both require innovations in device technology before they can be used in practical applications.

Introduction

This paper discusses research programs sponsored by NASA Lewis Research Center for work on optical computers. The objective of this research is to determine the role optical computers may play in future aircraft engine systems. Fiber optics are attractive for engine applications because of optics' inherent immunity to electromagnetic interference (EMI) and because optical signals can be safely transmitted through areas that contain explosive materials. A program was undertaken to develop optical sensors for engine control. Optical links were considered to be superior for signal transmissions aboard the aircraft, especially for off-engine mounted control computers. Currently, engine control computers are mounted on the engine. Special environmental packages must be designed to protect the computer from the harsh temperature and vibration environment. Off-engine mounted computers would be located in a more benign environment. Along with the optical sensors and optical transmission lines, optical computers are considered in this future technology program to determine if and how optical computers could be used. Optical computers are considered because of potential speed improvements over conventional processors. Future aircraft computers will be faster and will have a higher capacity to handle the more complex engines of tomorrow. Some of the functions onboard computers might be called on to perform, if the computers could operate in real-time and had the required capacity, are adaptive control, engine condition monitoring, and sensor failure accommodation. For adaptive control the computer must perform vector/matrix and matrix/matrix calculations. These operations are required to calculate optimal controller gains for multivariable control.

Engine condition monitoring requires an accurate nonlinear simulation of the engine, and this simulation must operate in real-time. Sensor failure accommodation also requires a nonlinear model to provide accurate state estimates of failed sensor states. Real-time simulation of nonlinear systems is, therefore, important for safe, efficient operation.

The optical computer has the potential for high speed operation required for adaptive control and has the potential for handling nonlinear simulations in real-time. Complex interfacing between optical sensors and the optical computer may be avoided if the optical sensor outputs can be directly processed by the optical computer. NASA Lewis Research Center is supporting two grants for basic research into optical computing and how optical computing could be used in aeronautical applications. Carnegie-Mellon University is developing a hybrid electro-optic computer capable of performing matrix operations such as inversion, multiplication, etc. Ohio State University is developing a digital optical processor. A liquid crystal light valve (LCLV) is used as the central computing element. With the LCLV, logic operations and arithmetic operations can be done.

Hybrid Electro-Optical Processor

Figure 1 illustrates possible steps used in calculating the optimal controller gains for adaptive control. From sensed information about the system to be controlled, a linearized model is developed using various identification techniques. Given a performance index, optimum control gains are calculated for the multivariable system by solving the Matrix Riccati Equation. At present this procedure is carried out on the ground. It would be desirable to do this onboard since the optimum gains could be updated to reflect changes in engine characteristics or variability from engine to engine. To do this onboard would require fast solutions to the Matrix Riccati Equation. Typically, a solution of the Matrix Riccati Equation takes hundreds of msec on an electronic computer. Using an optical computer to perform the matrix operations, solution times of a msec or better are possible.

The Carnegie-Mellon optical vector/matrix multiplier is shown in figure 2. The values of the components of the vector x are represented by the intensities of the light emitting diodes (LED) or Laser Diodes (LD). The LED outputs are connected to the matrix mask "A" with optical fibers. Variable transmittance of the cells of the "A" matrix represent the values of the elements a_{ij}. Because the A matrix can be bipolar, scaling and biasing to accommodate bipolar values must be done. These scaled, biased values are the values of the a_{ij} elements. The product of the intensity and mask cell transmittance are summed and focused on a linear detector array. This operation is performed rapidly on the optical computer because all operations are done in parallel. When combined with a microprocessor, the optical matrix/vector multiplier can be used to iteratively solve equations of the form $Ax = y$ where x is unknown (figure 3). A guess is made for the x vector. A new x value is calculated, and when the last value of x and the new value of x are equal within a prescribed tolerance, the solution converges and the implicit solution $A^{-1}y = x$ is obtained. The steady-state Matrix Riccati equation can be reduced to the form $Ax = y$ and solved by this method.

A photograph of the vector/matrix multiplier hardware used in lab demonstrations is shown in figure 4. The LED array shown is sealed with white RTV compound and bolted to the fiberoptic element, matrix mask and detector array. To have any practical application the mask used to represent the "A" matrix must be alterable to accommodate changing values of the elements a_{ij}. This concept does have the potential for rapid solutions of the Matrix Riccati Equation.

Digital Optical Processor

A second grant sponsored by NASA Lewis for work in optical computing is with Ohio State University. The objective of this work is to build and demonstrate optical analogs to digital components to perform various combinational and sequential logic functions. The main element used in this work is a liquid crystal light valve. Residue and binary arithmetic are both being considered for use by this optical computer.

The liquid crystal light valve shown in figure 5 operates in a controlled birefringence mode. Birefringence means that the refractive index is different when the polarization is in the direction of the principal axes compared to polarization at right angles to the axis. In the off state the cigar-shaped liquid crystal molecules are parallel to the glass face. This state has finite birefringence. This birefringence is decreased with application of the AC voltage. The molecules tip as a function of the excitation and light input.

Figure 6 illustrates how the LCLV operates. The electrical excitation can be adjusted to either of two values. Zero optical input produces either no change in the read beam or a rotation of 90° in the polarization of the read beam. The intensity of the input light beam is set so the complementary rotation is achieved when the beam is present.

The face of the liquid crystal light valve can operate on approximately 100 000 spots/in^2 (figure 7). Each spot is independent of its neighbors. This enables the LCLV to be used for operating on 2-D arrays of binary data. The use of residue arithmetic with this LCLV will make optimum use of the parallelism of the LCLV. This LCLV, with the extremely high density of spots, has potential for very large data transfer rates.

Ohio State has demonstrated a number of logic functions using the LCLV. Figure 8 illustrates an optical AND operation. The LCLV is biased to produce 90° rotation of polarization with bright input. As long as A and B spots both have bright inputs there is an output. Should either A or B go dark the output goes to zero.

Figure 9 shows an optical latch circuit. The Glan Thompson prism reflects vertically polarized light and transmits horizontally polarized light. The LCLV is biased to produce a 90° rotation of the reflected beam when the input to the back of the LCLV is bright. The latch is configured to load the data input when the clock is bright (high) and to store the input when the clock is dark (low). The stored output $\bar{Q}$ is the complement of the input; that is, when the input is bright, $\bar{Q}$ will be dark and vice versa. If both the clock and the input are bright, the input beam will be rotated 90° to vertical polarization and pass through the analyzer to supply a bright signal to the back of the LCLV. The horizontally polarized light from the DC light source will be rotated to vertical polarization upon reflection from the LCLV and will be reflected by the Glan Thompson prism, thus producing a dark output $\bar{Q}$. When the clock goes dark, the dark output $\bar{Q}$ will remain regardless of changes in the input. This occurs because the DC light source will continue to illuminate the back of the LCLV through the feedback loop. Similar reasoning can be used to explain the loading and storage of a dark input which produces a bright output $\bar{Q}$.

The LCLV used in these experiments is very slow, taking hundreds of millisecs to operate. For practical systems the spatial light modulator must operate in msecs or better to achieve the high data rates required for real-time computing.

Concluding Remarks

In summary, future aeronautical control systems will be more complex. Many functions not performed now and many functions done on the ground would better be done onboard if the speed and capacity of the computer were adequate. Among these functions are providing a real-time engine simulation for engine condition monitoring and for sensor failure accommodation. For adaptive control, onboard optimum multivariable control design would be done. For this to be done, rapid solutions of the Matrix Riccati Equation are required. Optical computers show some potential for applications to engine control systems because of their extremely high speed. Large data transfer capabilities and very fast solution of matrix equations are possible. However, much work has to be done in the area of faster spatial light modulators, optical read/write memories, and addressable alterable masks. These technology developments together with integrated optics may result in a practical electro-optic computer that could be used in engine control systems.

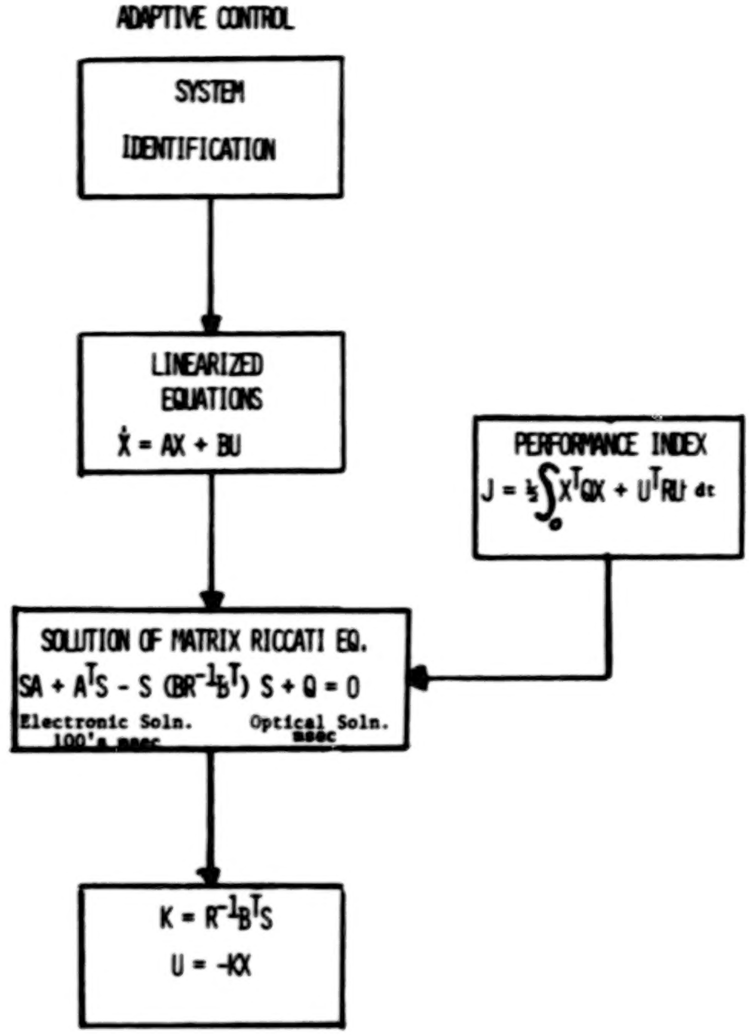

ADAPTIVE CONTROL
SYSTEM IDENTIFICATION
LINEARIZED EQUATIONS
$\dot{X} = AX + BU$
PERFORMANCE INDEX
$J = \frac{1}{2}\int_0 X^T QX + U^T RU \, dt$
SOLUTION OF MATRIX RICCATI EQ.
$SA + A^T S - S (BR^{-1}B^T) S + Q = 0$
Electronic Soln. 100's msec
Optical Soln. msec
$K = R^{-1}B^T S$
$U = -KX$

Figure 1.- Sequence of steps in adaptive control application.

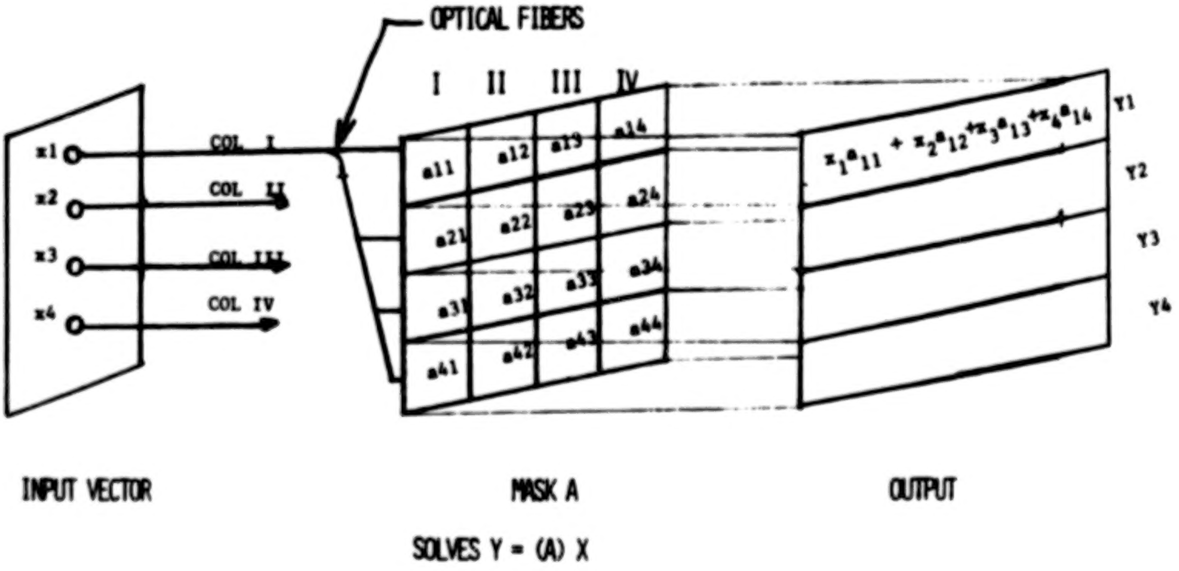

OPTICAL FIBERS
I
II
III
IV
x1
x2
x3
x4
COL I
COL II
COL III
COL IV
a11
a12
a13
a14
a21
a22
a23
a24
a31
a32
a33
a34
a41
a42
a43
a44
$x_1a_{11} + x_2a_{12} + x_3a_{13} + x_4a_{14}$
Y1
Y2
Y3
Y4
INPUT VECTOR
MASK A
OUTPUT
SOLVES Y = (A) X

Figure 2.- Vector matrix multiplier (analog model).

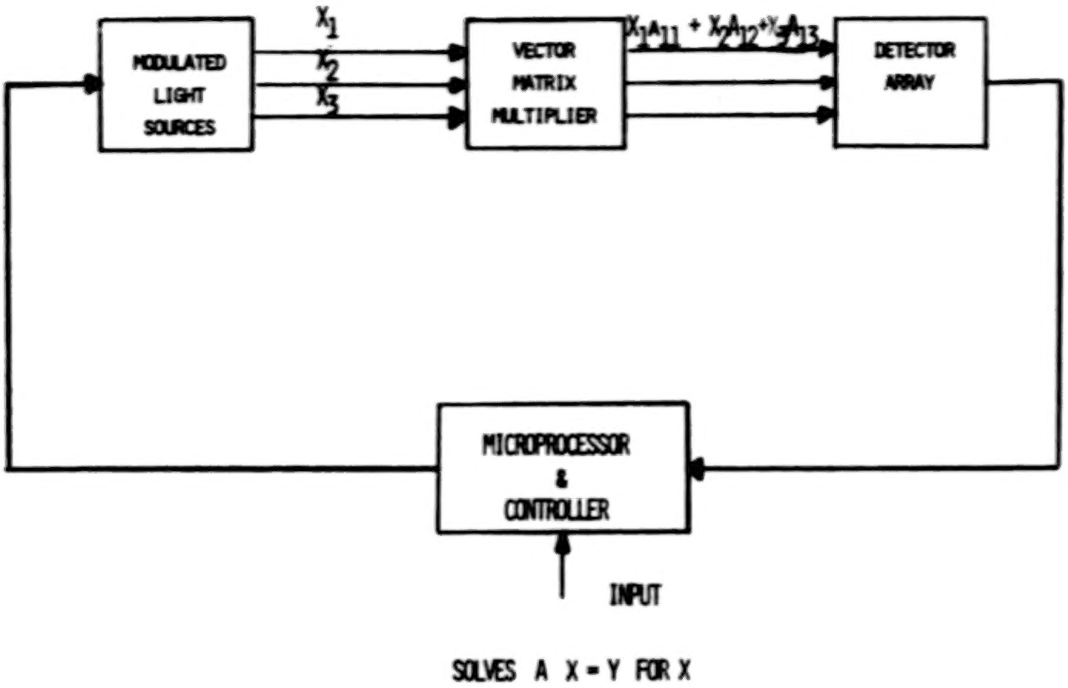

MODULATED LIGHT SOURCES
X1
X2
X3
VECTOR MATRIX MULTIPLIER
X1A11 + X2A12 + X3A13
DETECTOR ARRAY
MICROPROCESSOR & CONTROLLER
INPUT
SOLVES A X = Y FOR X

Figure 3.- Solution of vector/matrix equation using iterative optical processor.

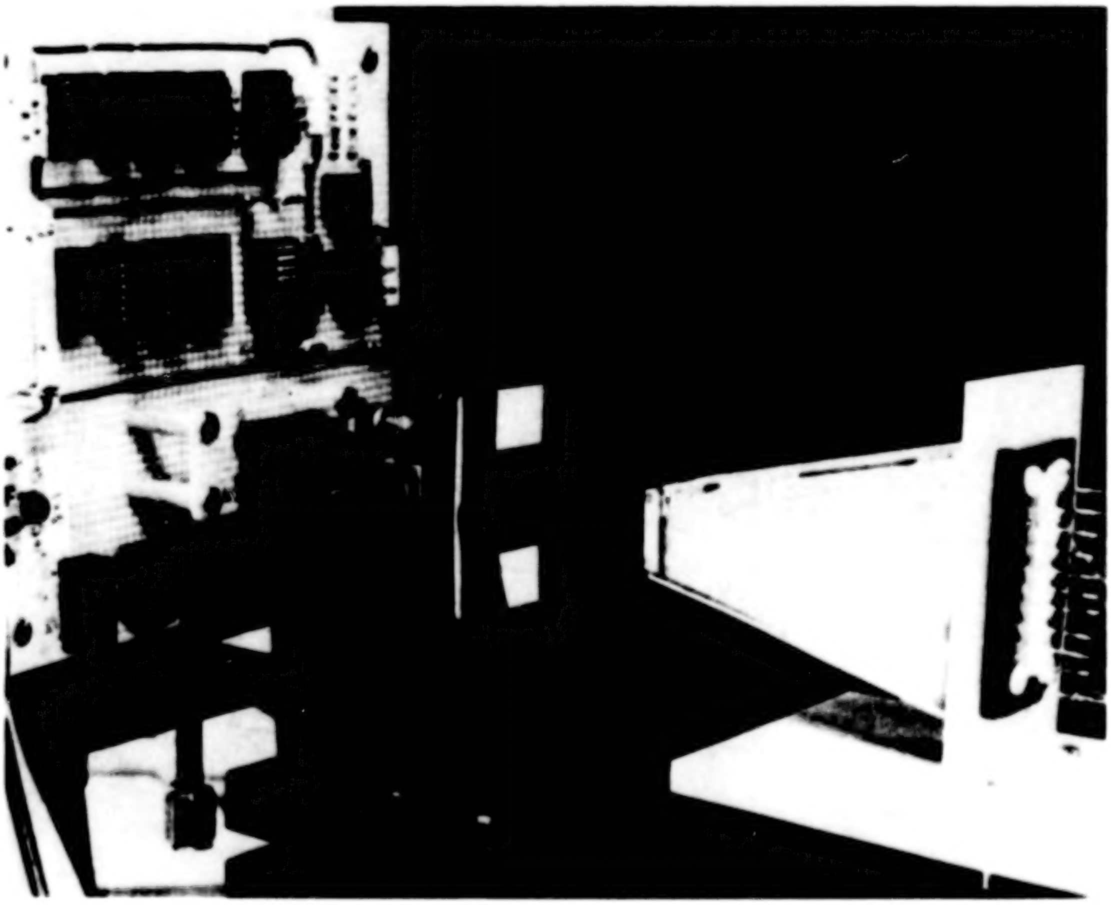


Figure 4.- Iterative optical processor hardware.

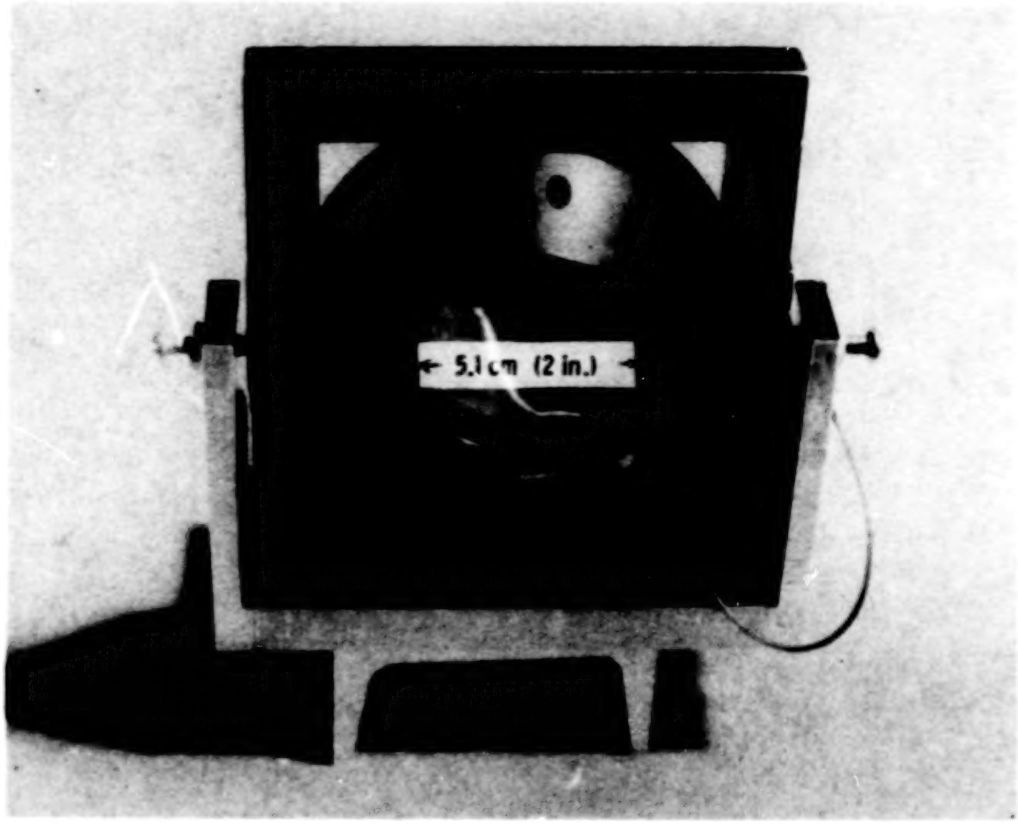
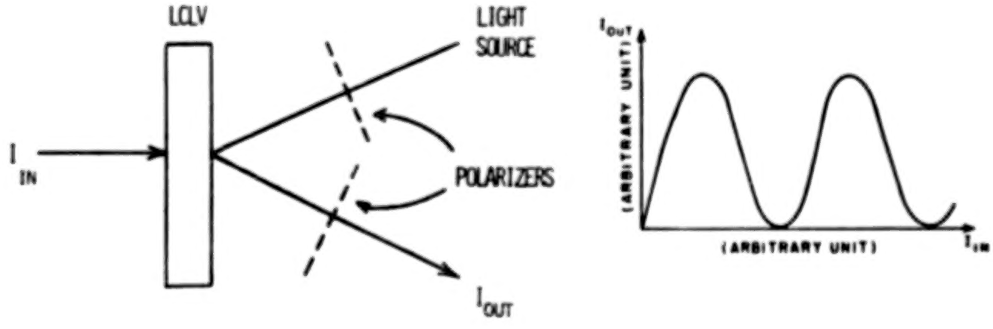


Figure 5.- Liquid crystal light valve.

HUGHES LIQUID CRYSTAL LIGHT VALVE (LCLV)
SPATIAL LIGHT MODULATOR (SLM)


LCLV
LIGHT SOURCE
I_{IN}
POLARIZERS
I_{OUT}
I_{OUT}
(ARBITRARY UNIT)
(ARBITRARY UNIT)
I_{IN}

- THE LCLV ACTS LIKE A MIRROR WHICH ROTATES THE POLARIZATION OF A REFLECTED BEAM
- THE AMOUNT OF ROTATION DEPENDS ON I_{IN}
- THE LCLV CAN OPERATE INDEPENDENTLY ON THOUSANDS OF SEPARATE BEAMS (SPOTS)

Figure 6.- Operation of liquid crystal light valve.

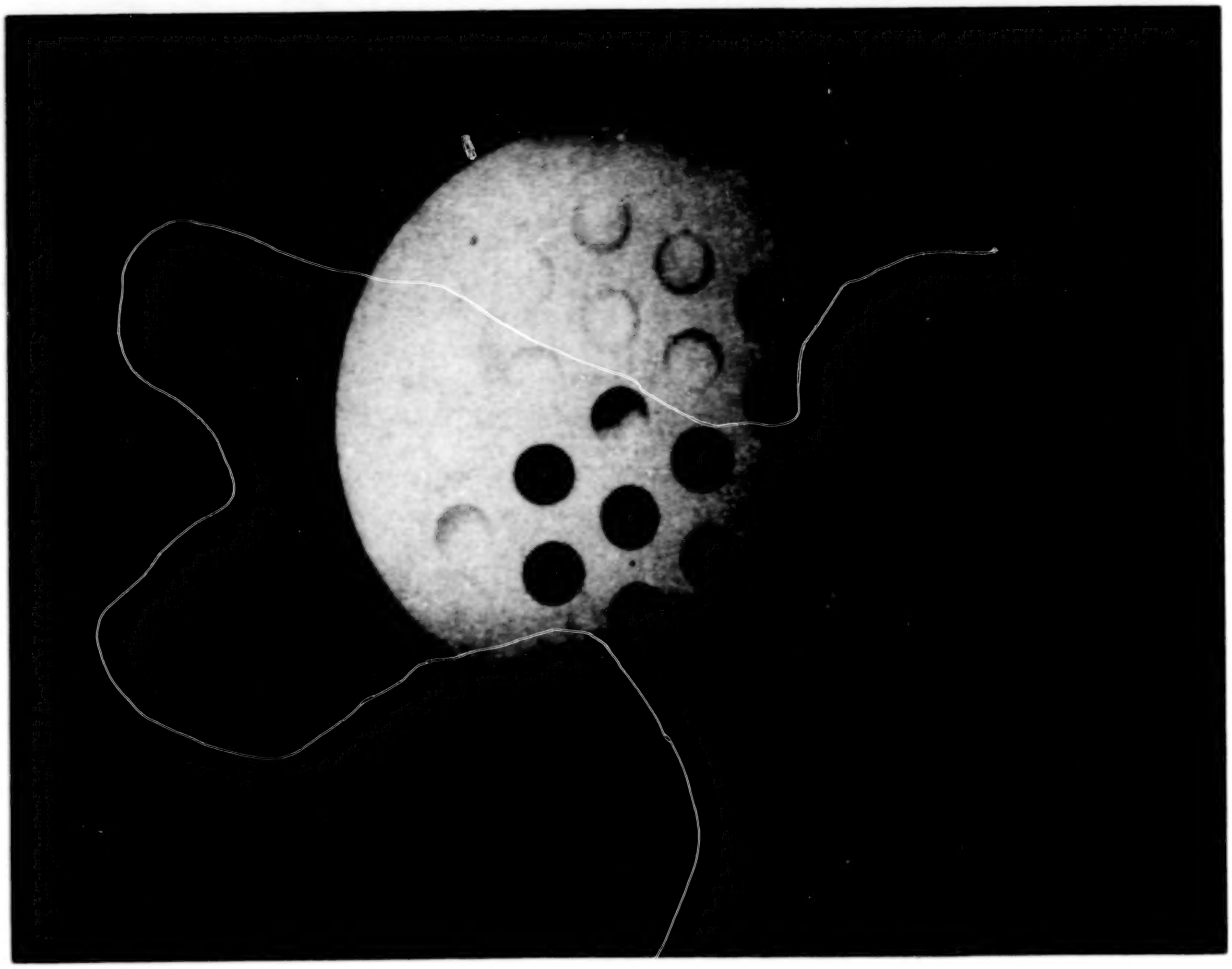


Figure 7.- Liquid crystal light valve face showing individual cells.

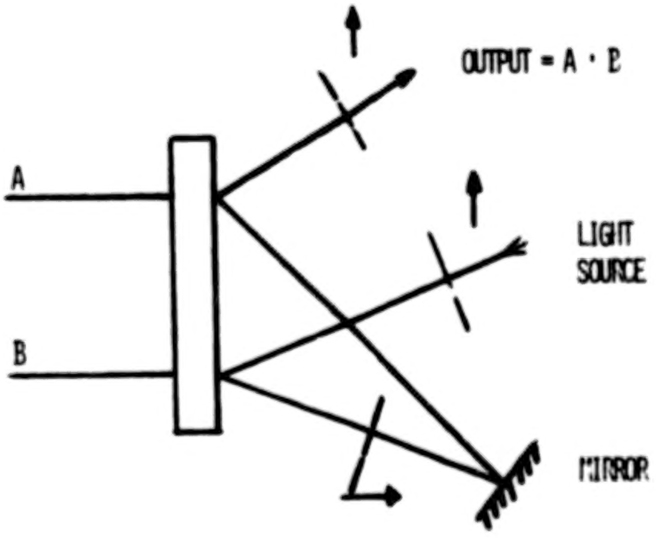

OUTPUT = A · B
A
B
LIGHT SOURCE
MIRROR

Figure 8.- Logical **AND** operation using liquid crystal light valve.

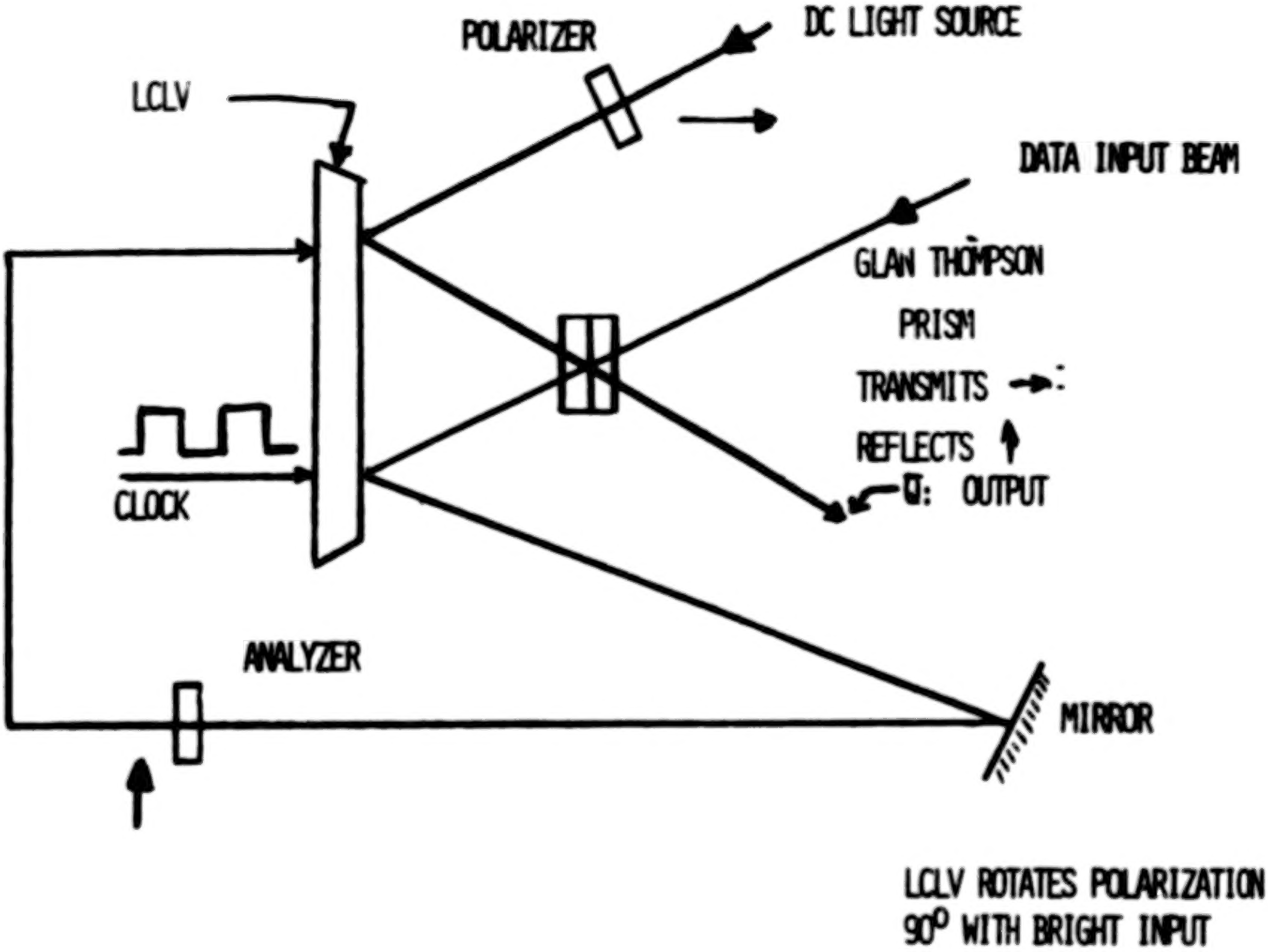

POLARIZER
DC LIGHT SOURCE
LCLV
DATA INPUT BEAM
GLAN THOMPSON
PRISM
TRANSMITS →
REFLECTS ↑
Q: OUTPUT
CLOCK
ANALYZER
MIRROR
LCLV ROTATES POLARIZATION
90° WITH BRIGHT INPUT

Figure 9.- Optical latch circuit.

OPTICAL DIGITAL COMPUTERS*

Alan Huang
Bell Laboratories
Holmdel, NJ 07733

Advances in integrated optics, optical communications, optical bistability, and optical disk storage have prompted a renewed interest in studying the possibility of an optical digital computer. The limitations in utilizing optics stem from the fact that it is very difficult to get one optical signal to affect another. This makes it difficult to perform switching. One approach around this problem is to detect one optical signal with a pn junction and stimulate another signal with another pn junction. The need for pn junctions immediately makes the use of optics questionable. The potential of optics ironically stems from the same inability for interaction. Lenses, prisms, and mirrors can easily communicate thousands of channels of information in parallel. The relevance of this to computers is that the throughput of future systems will be limited by clock skew, bandwidth, interconnection complexity, pin limitations, and the Von Neumann bottleneck. These are all communications rather than switching problems. Unfortunately, there is no easy way to incorporate the communications capabilities of optics in conventional processors since the architecture of these processors have been adapted by evolution to the implicit tradeoffs of an electronics technology. Such a change is so fundamental that a careful review of the basic structure and function of conventional processors is necessary.

---

* Because this paper was not available at the time of publication, only the abstract is included.

A design for a optical digital computer is presented to illustrate how some of the unique properties of optics such as its innate parallelism and non-interfering propagation can be used to overcome some of the limitations of future computers. The basic approach involves decomposing the structure of a finite state machine, the distant ancestor of current computers, into a logic unit, an interconnection array, and a latching unit; and then implementing each of these units with optics. The logic unit for such a processor can be constructed from an optical NOR gate array. Such a logic array can be implemented by projecting several binary images on a common surface, inverting, and thresholding the intensity of the result. Each pixel functions as a NOR gate and such a NOR operation is sufficient to establish a complete logical set. The latching unit which serves as the memory for the finite state machine can be implemented with various bistable optical devices. The interconnection array can be implemented with a hologram or conventional optics which images each pixel(x, y) to pixel(x, y+1), pixel(x+1, y+1), and pixel(x-2,y+1). This interconnection pattern establishes a complete connective set which can be customized to provide arbitrary interconnections.

This basic architecture, while abstract and quite primative, can be extended, specialized, and applied to a wide range of problems in image and parallel processing.

HIGH SPEED TECHNIQUES FOR SYNTHETIC APERTURE RADAR IMAGE FORMATION*

Demetri Psaltis
California Institute of Technology
Pasadena, CA 91125

Armand R. Tanguay, Jr.
University of Southern California
Los Angeles, CA 90007

and

Thomas J. Bicknell
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109

A synthetic aperture radar system produces data requiring rather complex processing to form the image of the terrain being mapped. The process is essentially a two-dimensional convolution which is usually separated into two one-dimensional convolutions with spatial variance of the reference function in one dimension. The reference functions may be quite large (1000 elements or more) requiring high computation rates for digital systems. Optical systems have a significant advantage due to the parallelism of one- or two-dimensional processing. The linear FM variation of the reference function in both dimensions makes the optical system rather simple. Since the point target phase histories are self-focusing, like Fresnel zone plates, only bandpass filtering is required for moderate resolution systems. The spatial variance of the reference function or phase history in the along-tract dimension is linear and can be compensated for by tilting either the input and

---

* Because this paper was not available at the time of publication, only the abstract is included.

output image planes or a cylindrical lens. An additional complication of higher resolution systems is the requirement to compensate for range migration. This correction is usually performed by lenses used as a phase filter in the transform plane. A conventional radar correlator is shown in figure 1. The range telescope images both the range and the azimuth focal planes. The azimuth telescope affects only the azimuth or along-track dimension, bringing it into coincident focus with range.

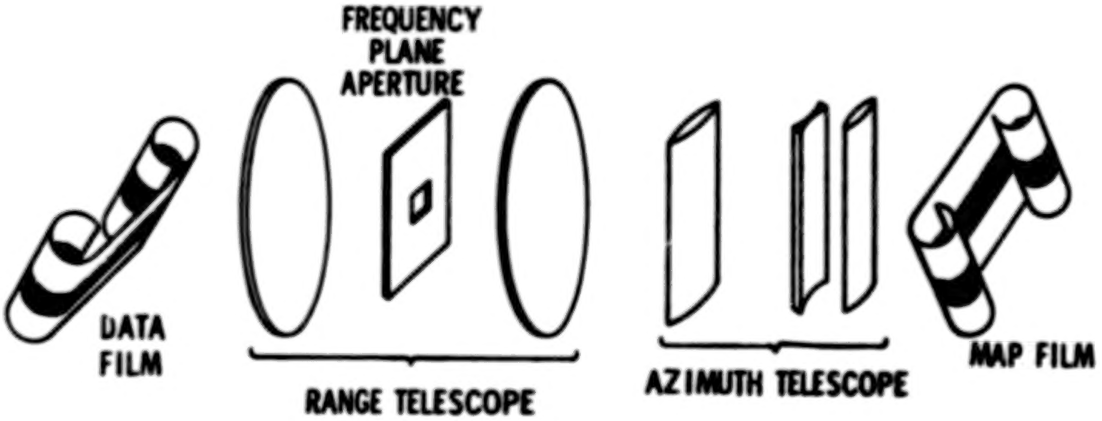


Figure 1.- Film input coherent synthetic aperture radar processor.

The desire to achieve real time throughput, higher reliability, and greater accuracy has led to the development of high speed digital systems for SAR data processing. However, such systems are quite large, costly, and power consumptive, even using recent and expected future developments in digital technology. Many radar applications, such as quick-look or search modes, require or would be better served by onboard processing. These systems require small, lightweight processors, especially for use in conjunction with spacecraft-borne radars. A good example is the upcoming VOIR mission where onboard processing would allow transmission of 10 times as much data as the currently planned ground processing scheme allows.

One possible approach to high speed synthetic aperture radar signal reconstruction involves the utilization of two-dimensional real time spatial light modulators as recyclable replacements for photographic film in the input transducer plane of a modified SAR-coherent optical processor, as shown in figure 2. Leading candidate spatial light modulators include modified Pockels readout optical modulators (PROM), CCD-addressed liquid crystal light valves, and CCD-addressed membrane light modulators. The fundamental physical limitations affecting SAR-processor performance characteristics of such real time devices are under investigation. Current research on the PROM is focused on the effects of device operational mode, device constitutive parameters, electro-optic crystal orientation, writing wavelength, frame rate/data overwrite/presuming, erasure completeness, and image retention on the overall quality of SAR image formation. Both modulated laser scanning and intensified CRT temporal-to-spatial input approaches are being examined.

A spatial light modulator (SLM) that is commonly used in one-dimensional optical processors is the acousto-optic delay line. We are investigating the possibility of using acousto-optic devices (AOD's) as the input SLM in an optical SAR processor. AOD's cannot be used in a conventional SAR processor, however, since they can only introduce one-dimensional spatial modulation in the optical system. In the architecture we are investigating, the two-dimensional processing is accomplished utilizing one spatial dimension and one time dimension. Hence, a one-dimensional spatial modulator can be used. A schematic diagram of the processor is shown in figure 3. The backscattered signal for each transmitted pulse is applied to the AOD. The system is illuminated with a pulse light source in order to "freeze" the traveling acoustic wave in the AOD. The range compression operation is accomplished by

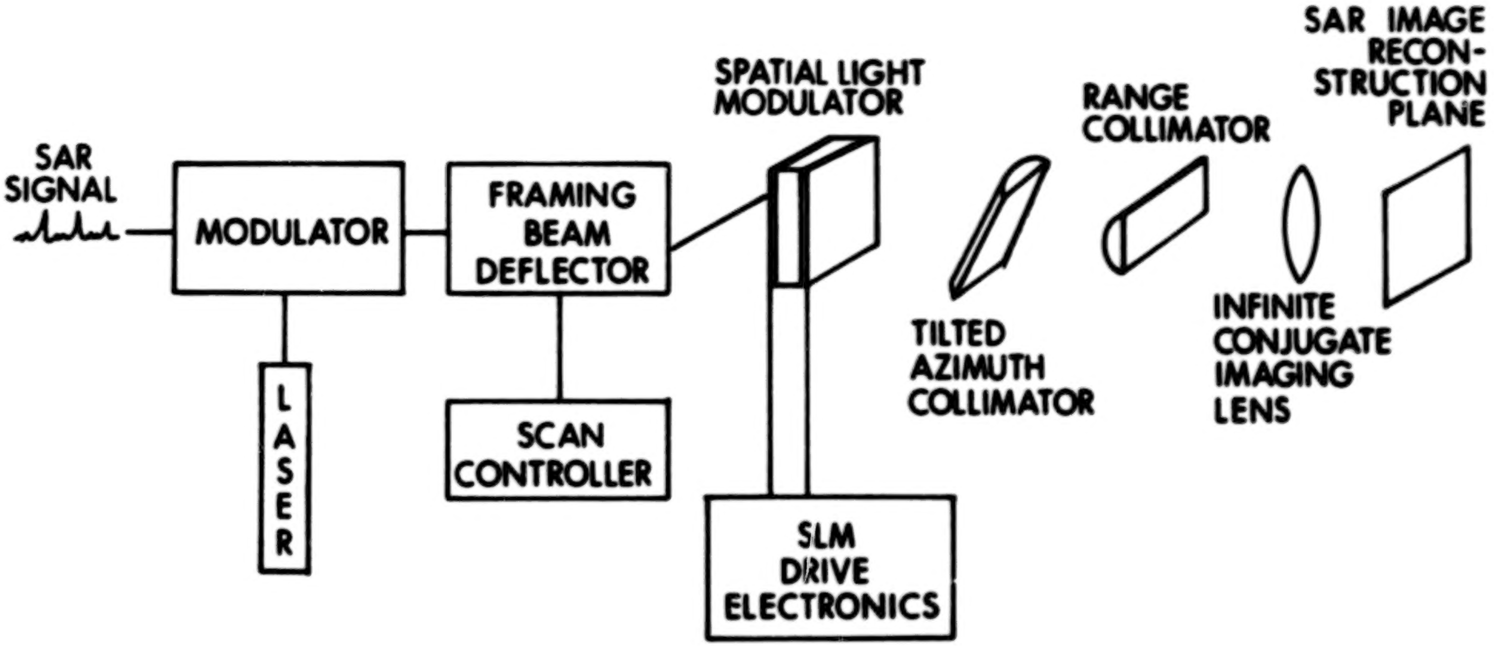

SAR SIGNAL
MODULATOR
LASER
FRAMING BEAM DEFLECTOR
SCAN CONTROLLER
SPATIAL LIGHT MODULATOR
SLM DRIVE ELECTRONICS
TILTED AZIMUTH COLLIMATOR
RANGE COLLIMATOR
INFINITE CONJUGATE IMAGING LENS
SAR IMAGE RECONSTRUCTION PLANE

Figure 2.- Real time coherent synthetic aperture radar processor utilizing a two-dimensional spatial light modulator.

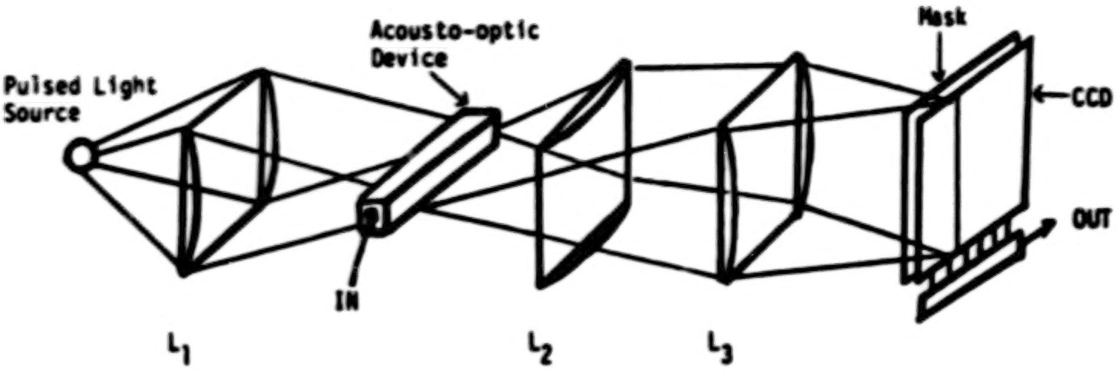


Figure 3.- Real time coherent synthetic aperture radar processor utilizing a one-dimensional acousto-optic modulator and a CCD-array detector.

the focusing action of lens L_2. The photogenerated charge pattern stored in the CCD is shifted vertically by one pixel after each radar pulse. The signal that reaches the output stage of the CCD at each horizontal (range) location is the correlation of the range compressed signal and the transmittance of the mask placed immediately in front of the CCD detector. Azimuth compression is performed by making the transmittance of the mask a linear FM function in the vertical direction.

Utilization of a two-dimensional spatial light modulator in a real time SAR processor allows for simplicity in image formation, and is, in addition, inherently flexible for handling different radar signal types. A disadvantage in this approach is the relatively recent development status of candidate spatial light modulators. The state-of-the-art of acousto-optic devices is, on the other hand, further advanced and represents a relatively mature technology. The limitations in the AOD/CCD approach are expected to be the dynamic range and number of elements of the CCD detector array. It is hoped that such optical signals in an integrated, compact, and rugged format. Many integrated real time SAR processors will initiate new mission applications not presently envisioned due to current digital processor limitations.

OPTICAL IMPLEMENTATION OF SYSTOLIC ARRAY PROCESSING

H. J. Caulfield*, W. T. Rhodes+, M. J. Foster†, and Sam Horvitz-

- * Innovative Optics, Inc., P.O. Box 1275, Concord, MA. 01742
- + Georgia Institute of Technology, School of Electrical Engineering, Atlanta, GA. 30332
- † Carnegie-Mellon University, Computer Science Department, Pittsburgh, PA. 15213
- - Naval Underwater Systems Center, PA4, New London, CT. 06320

SUMMARY

We show how certain algorithms for matrix-vector multiplication can be implemented using acousto-optic cells for multiplication and input data transfer and using CCD detector arrays for accumulation and output of the results. No 2-D matrix mask is required; matrix changes are implemented electronically. A system for multiplying a 50-component nonnegative-real vector by a 50 x 50 nonnegative-real matrix is described. Modifications for bipolar-real and complex-valued processing are possible, as are extensions to matrix-matrix multiplication and multiplication of a vector by multiple matrices.

INTRODUCTION

During the past several years, Kung and Leiserson at Carnegie-Mellon University (refs. 1,2) have developed a new type of computational architecture which they call "systolic array processing". Although there are numerous architectures for systolic array processing, a general feature is a flow of data through similar or identical arithmetic or logic units where fixed operations, such as multiples and adds, are performed. The data tend to flow in a pulsating manner, hence the name "systolic". Systolic array processors appear to offer certain design and speed advantages for VLSI implementation over previous calculational algorithms for such operations as matrix-vector multiplication, matrix-matrix multiplication, pattern recognition in context, and digital filtering. This paper grew out of our desire to explore the possibility of improving systolic array processors by using optical input and output. We will concentrate on describing the particular case of matrix-vector multiplication, but note that many other operations can be performed in an analogous manner.

SYSTOLIC MULTIPLICATION OF A VECTOR BY A MATRIX

The problem we address is that of evaluating a vector $\vec{y}$ given by

$$\vec{y} = \underline{A}\vec{x} \tag{1}$$

where $\underline{A}$ is an n by n matrix, and $\vec{x}$ and $\vec{y}$ are n-component vectors. We assume that $\underline{A}$ has bandwidth w, i.e., all of its non-zero entries are clustered in a band of width w around the major diagonal. Such matrices arise frequently in the solution of boundary value problems for ordinary differential equations. A systolic array that solves this problem is introduced by Kung and Leiserson (ref. 1,2) and will be reviewed briefly here.

A systolic array for multiplying a matrix of bandwidth w by a vector of arbitrary length has inner-product cells. The array for bandwidth 4 is shown in figure 1. Each of the four heavy boxes represents an inner-product cell, capable of updating the vector component y_i according to the replacement

$$y_i \leftarrow y_i + a_{ij}x_j \tag{2}$$

The cells act together at discrete time intervals, or beats, with half of the cells active on each beat. The elements of the matrix $\underline{A}$ are input from the right, and the vector $\vec{x}$ is input from the top. Zeroes are input from the bottom, and accumulate terms of the vector $\vec{y}$ as they move upward.

Figure 2 traces the action of the array for several beats, or pulsations, showing the terms of $\underline{A}$ and $\vec{x}$ and the partial terms of $\vec{y}$ that are in each cell on each pulsation. Thus on pulsation 1, $y_1 = 0$ is entered. In pulsation 2, x_1 is entered. In pulsation 3, y_1 becomes $a_{11}x_1$. In pulsation 4, y_1 becomes $a_{11}x_1 + a_{12}x_2$. In pulsation 5, y_1 exits. Every other pulse another y_j exits and on that same pulse another y_k is inserted (at an initial value of zero).

OPTICAL SYSTOLIC ARRAY PROCESSING

Key features of the systolic array approach to matrix-vector multiplication are (1) a regular, directed flow of data streams, (2) multiplication, and (3) addition or accumulation. These features are also characteristic of many optical signal processing systems, and it should come as no great surprise that optical implementations of systolic architectures are possible. Since both bulk and surface acoustic waves are routinely used in optical signal processing to produce a moving stream of data and for multiplication of data, it seems natural to use these components for optical systolic array processing.

We choose as our example the simple matrix-vector multiplication

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{3}$$

assuming initially that all quantities in this equation are real and nonnegative. The basic concept is illustrated with the help of figure 3. The system shown consists of an acoustooptic modulator illuminated by the collimated light from three LEDS, a Schlieren imaging system, and three detectors connected to a CCD analog shift register. At the moment illustrated in the figure, modulating signals proportional to x_1 and x_2 have been input to the acousto-optic modulator driver, producing short grating segments in the acousto-optic cell. As the x_1 grating segment passes in front of LED number 2 (the situation shown in the figure), that LED is pulsed in proportion to matrix coefficient a_{11}. The transmitted light, proportional in intensity to $a_{11}x_1$, is imaged onto CCD detector 2, which sends a proportional charge to an associated "bin" in the shift register.

The x_1 and x_2 grating segments now travel so as to be in front of LED's 1 and 3, respectively. At the same time, the accumulated CCD charge from detector 2 is shifted one bin, in the direction indicated by the arrow labeled "output" in the figure. LED's 1 and 3 are now pulsed in proportion to a_{21} and a_{12}, respectively. Since these LED's illuminate detectors 3 and 1 via grating segments x_1 and x_2, charge is generated by these detectors in proportion to $a_{21}x_1$ and $a_{12}x_2$, respectively, and accumulated in the corresponding shift register bins.

In the next increment of the system, charges are again shifted, with accumulated charge in proportion to $a_{11}x_1 + a_{12}x_2$, or y_1 being output. The charge packet now associated with detector 2 (already proportional to $a_{21}x_1$) is augmented by a final strobe of LED 2 by an amount proportional to $a_{22}x_2$. A final two shifts of the CCD charge packets bring charge proportional to $a_{21}x_1 + a_{22}x_2$, or y_2, to the output, and the operation is complete.

The system illustrated is easily expanded to accommodate matrix-vector operations of higher dimensionality. If $\vec{y}$ and $\vec{x}$ are N-component vectors and $\underline{A}$ an N x N matrix, the maximum number of LED's required is 2N-1 (the number of diagonals of the matrix), and the number can be smaller if $\underline{A}$ has a smaller bandwidth.

Numerous variations of the system of figure 3 are possible. Figure 4, for example, shows the LED's replaced by a single light source and an array of modulators. The CCD shift register has been replaced by stationary detectors and integrators combined with a second acousto-optic cell, which serves to deflect light to the correct detector/integrator. The acousto-optic deflector approach to sorting output data may facilitate greater system dynamic range than is achievable with CCD detector arrays.

BIPOLAR AND COMPLEX-VALUED COMPUTATIONS

It was assumed in the preceding section that all elements of the matrix and input vectors were nonnegative-real. In practice, most matrix-vector multiplication operations of importance involve bipolar-real or complex-valued vectors and matrices, and some means must be employed for handling them. If the elements are real valued, but not necessarily nonnegative, a two-component decomposition scheme described in ref. 3 can be employed. For complex-valued processing, several schemes have been described (ref. 4). One of these involves a three-component decomposition of complex numbers according to ref. 5,

$$z = z_0 + z_1 \exp\left[i2\pi/3\right] + z_2 \exp\left[i4\pi/3\right] \quad , \tag{4}$$

where z_0, z_1 and z_2 are nonnegative-real. Another involves biased real and imaginary components (ref. 6). All such methods lead to some additional processor complexity and to a reduction in the size of the vectors and matrices that can be accommodated.

OPERATING PARAMETERS OF A TYPICAL SYSTEM

Matrix size limitations are imposed by the acousto-optic modulator. Consider a system using for input a bulk acousto-optic cell with a 100 MHz bandwidth and a 10 μsec time window. We estimate that such a cell should accommodate 100 LED/lenslet combinations operating side by side, allowing multiplication of a 50-component nonnegative-real vector by a 50 x 50 nonnegative-real matrix. Achievable dynamic range depends on CCD detector dynamic range and on the correction of LED and acousto-optic modulator nonlinearities; it is too speculative to suggest numbers at this time. Operating speed is determined by the amount of time it takes to shift the components of $\vec{x}$ through the acousto-optic cell, plus setup and final readout time. For the 10 μsec window cell under consideration, it takes 5 μsec to get the x_1 grating segment to the middle of the acousto-optic cell, at which time the first LED pulse occurs. The last LED pulse occurs 10 μsec later, when x_{50} finally passes the midpoint of the cell. Following that pulse, an additional 50 μsec are required to read y_{50} out of the shift register. The time required for the 50 x 50 matrix-vector multiplication is thus 10 μsec processing time and 10 μsec latency, for a total of 20 μsec. During the processing interval, a total of 2500 multiplications are performed, at a rate of 2.5 x 10^8 multiplications per second. With suitable encoding of the data (refs. 3,4), this corresponds to a processing rate of 6.25 x 10^7 bipolar-real multiplications per second or 2.78 x 10^7 complex multiplications per second.

VARIATIONS

The system described does not exploit the two-dimensionality of the optical system. More than one matrix can multiply the same input vector at the same time if the single linear LED/lenslet and detector arrays are replaced with a collection of linear arrays, one above the other. Shear wave acousto-optic modulators, with nearly square window formats, can accommodate perhaps 20 such linear arrays, allowing 20 separate matrices to multiply the same input vector at the same time.

Matrix-matrix multiplication can be performed with related systems using multiple acousto-optic cells, or, alternatively, single cells with multiple driver/transducers. Figure 5 shows one possible arrangement for multiplication of two 2 x 2 nonnegative-real matrices. In general, for such a scheme, multiplication of two N x N matrices requires two multi-transducer acousto-optic modulators with 2N-1 transducers each. Alternatively, one such multi-transducer cell could be used, illuminated by a 2-D array of N^3-2 LED's.

REFERENCES

1. H. T. Kung and C. E. Leiserson, "Systolic Array Apparatuses For Matrix Computations", U.S. Patent Application, Filed December 11, 1978.

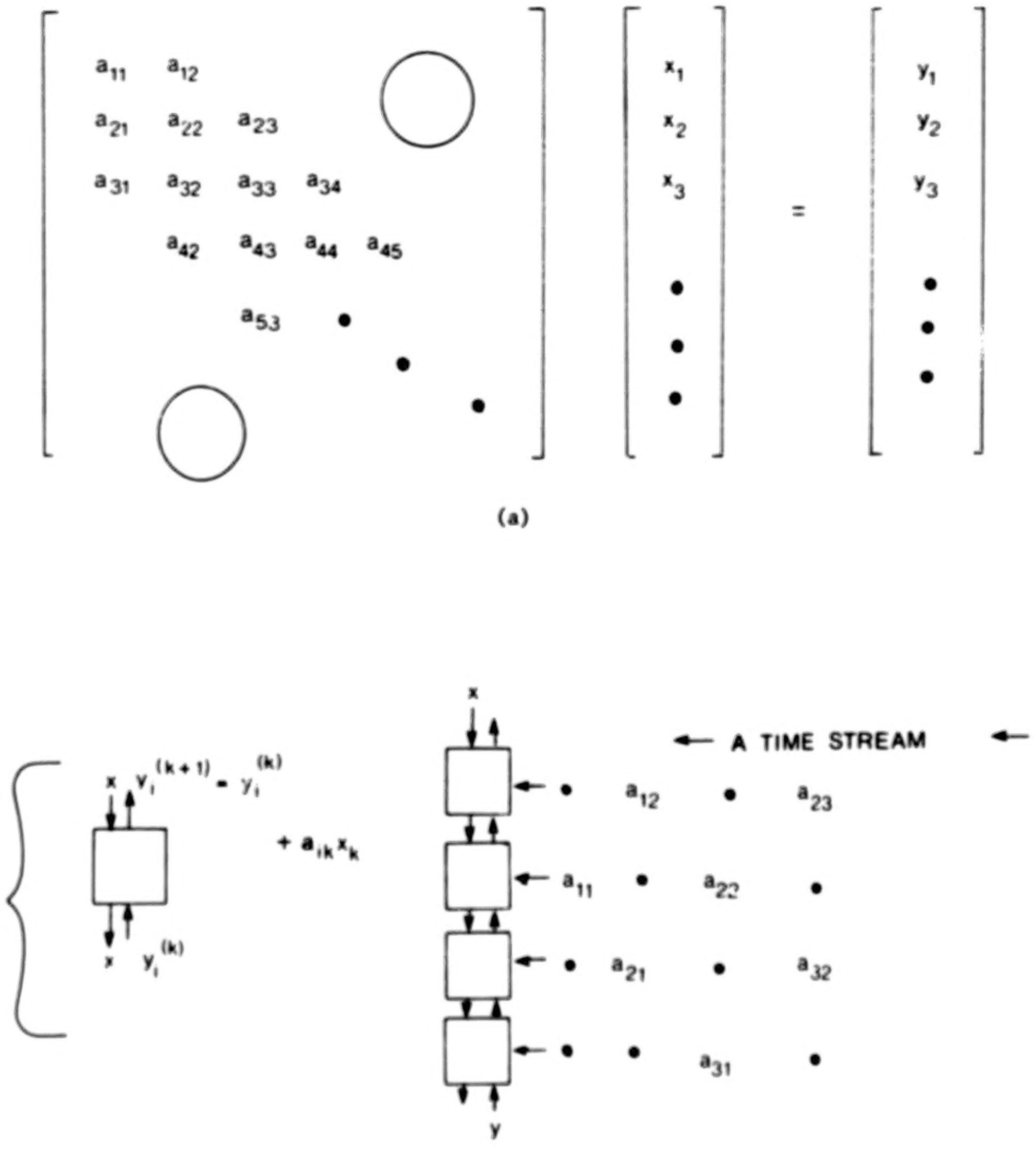
2. H. T. Kung and C. E. Leiserson, "Systolic Arrays for VLSI", in Introduction To VLSI, by C. A. Mead and L. A. Conway (Addison-Wesley, Reading, Mass., 1980).

3. H. J. Caulfield, D. Dvore, J. W. Goodman and W. T. Rhodes, "Eigenvector Determination By Noncoherent Optical Methods", Appl. Opt. 20, (1981) 2263-2265.

4. A. R. Dias, "Incoherent Optical Matrix-Vector Multiplication For High-Speed Data Processing", Ph.D Dissertation, Stanford University, 1980, (University Microfilm No. 8024641).

5. J. W. Goodman, A. R. Dias, and L. M. Woody, "Fully Parallel, High-Speed Incoherent Optical Method For Performing Discrete Fourier Transforms", Opt. Lett. 2, (1978) 1-3.

6. J. W. Goodman, A. R. Dias, L. M. Woody, and J. Erickson, "Some New Methods For Processing Electronic Image Data Using Incoherent Light", In Optica Hoy y Manana (Proceedings Of ICO-11 Conference, Madrid, Spain, 1978), J. Bescos, A. Hidalgo, L. Plaza, J. Santamaria, eds., pp. 139-145.


a11 a12
a21 a22 a23
a31 a32 a33 a34
a42 a43 a44 a45
a53
x1
x2
x3
y1
y2
y3
=
x
y_i^(k+1) = y_i^(k)
+ a_ik x_k
x
y_i^(k)
x
A TIME STREAM
a12
a23
a11
a22
a21
a32
a31
y

(a)

(b)

(c)

Figure 1.- Systolic multiplication of vector $\vec{x}$ by banded matrix $\underline{A}$. Traditional representation of this operation is shown in (a). Basic cell for this operation is shown in (b). Flow of $\vec{x}$, $\vec{y}$ and $\underline{A}$ data is shown in (c).

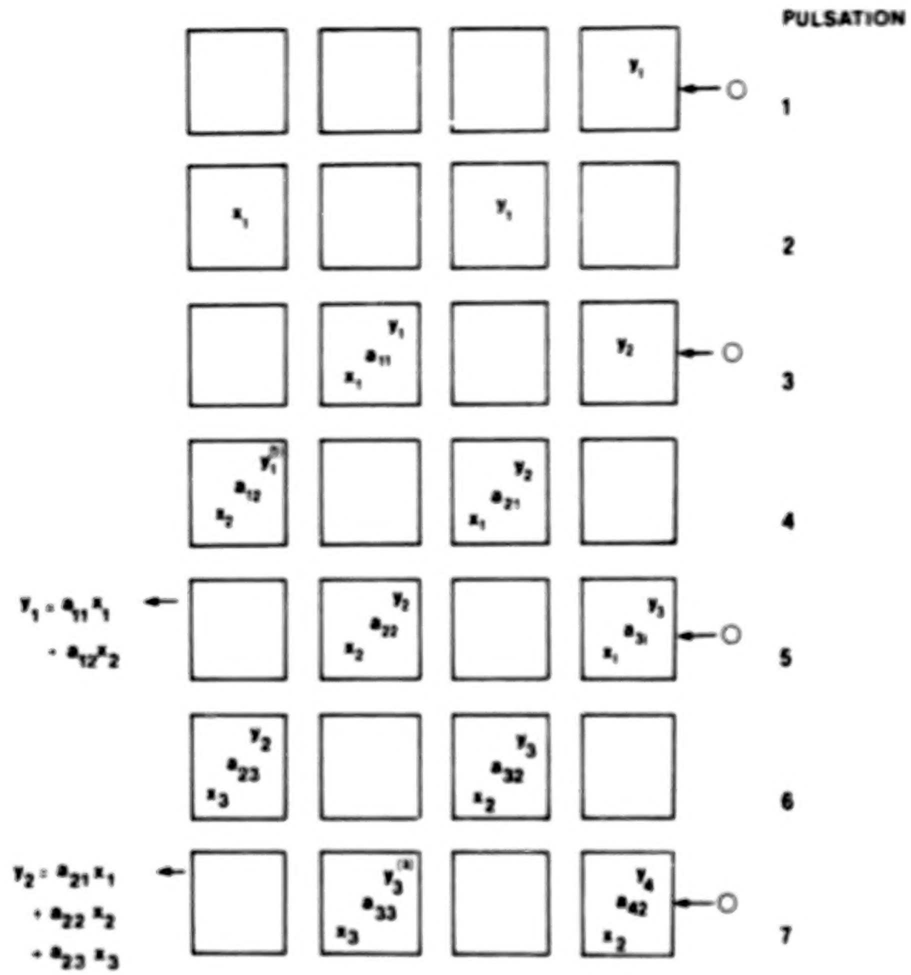

PULSATION
1
2
3
4
5
6
7
y_1
x_1
a_{11}
a_{12}
x_2
y_2
a_{21}
$y_1 = a_{11}x_1 + a_{12}x_2$
a_{22}
y_3
a_{31}
a_{23}
x_3
a_{32}
$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$
a_{33}
y_4
a_{42}

Figure 2.- The first seven pulsations of the processor of figure 1(c) are shown here and described in the text.

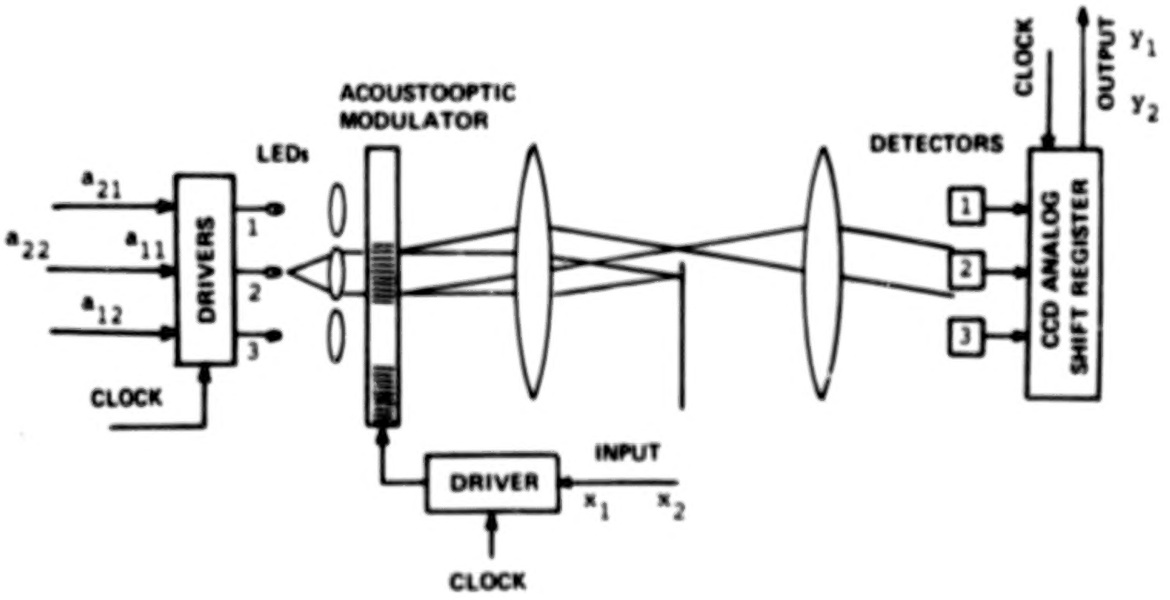

ACOUSTOOPTIC MODULATOR
LEDs
a_{21}
a_{22}
a_{11}
a_{12}
DRIVERS
1
2
3
CLOCK
DRIVER
INPUT
x_1
x_2
CLOCK
DETECTORS
1
2
3
CCD ANALOG SHIFT REGISTER
CLOCK
OUTPUT
y_1
y_2

Figure 3.- Optical implementation of systolic array processor of figure 1(c).

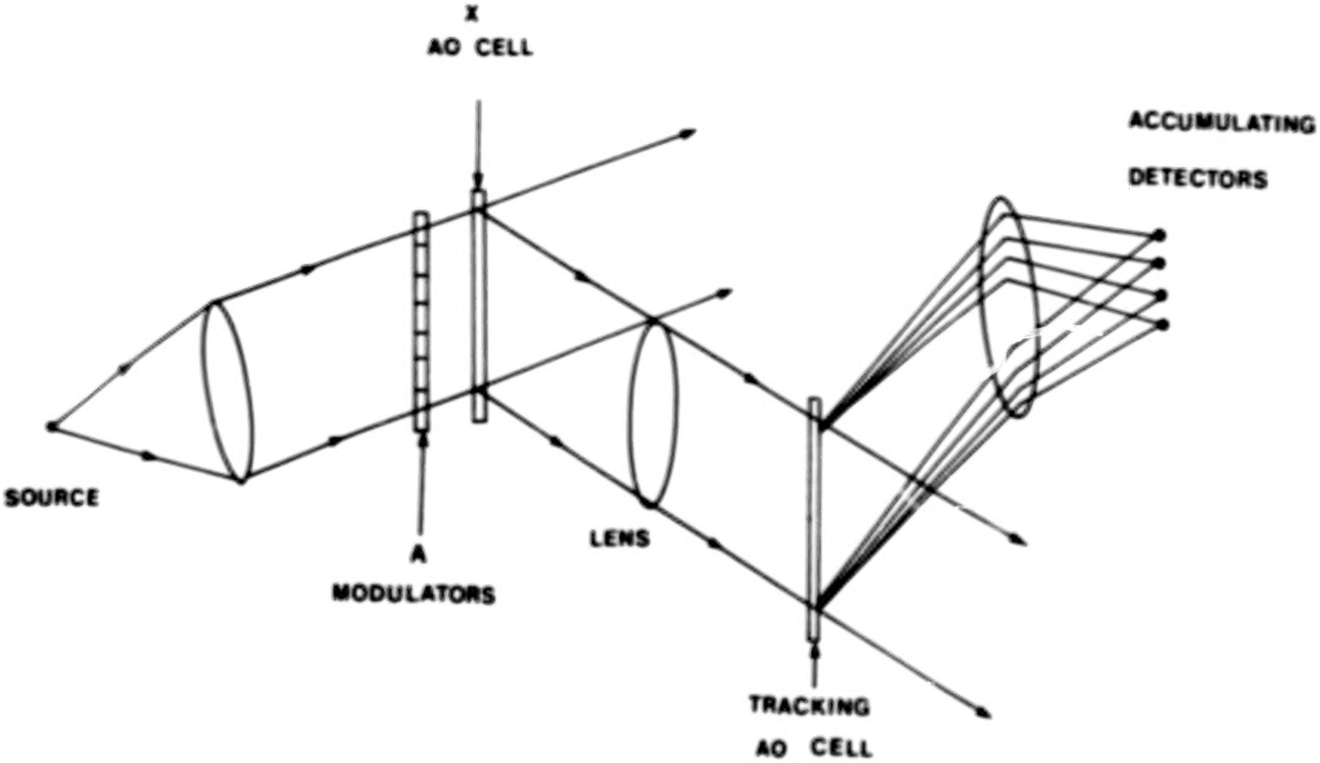
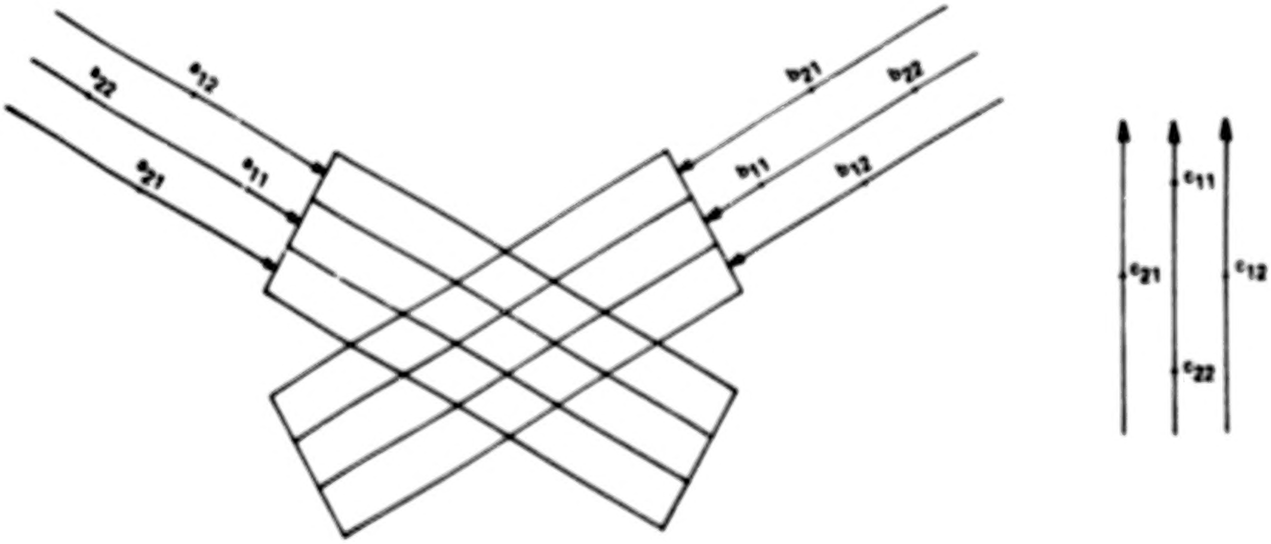

X
AO CELL
ACCUMULATING
DETECTORS
SOURCE
A
MODULATORS
LENS
TRACKING
AO CELL

Figure 4.- An alternative optical implementation of the processor of figure 1(c).


a_{12}
a_{22}
a_{21}
a_{11}
b_{21}
b_{22}
b_{11}
b_{12}
c_{11}
c_{21}
c_{12}
c_{22}

(a)

(b)

Figure 5.- Use of crossed acousto-optic cells to produce $\underline{A} \times \underline{B} = \underline{C}$. Input information flow is shown in (a), and calculated $\underline{C}$ values are produced as indicated in (b).

MATRIX VECTOR MULTIPLIER

WITH TIME VARYING SINGLE DIMENSIONAL SPATIAL LIGHT MODULATORS

A. Tarasevich and N. Zepkin
Lockheed Electronics Company
Plainfield, New Jersey

W. T. Rhodes
Georgia Institute of Technology
Atlanta, Georgia

ABSTRACT

Optical Matrix Vector Multipliers (OMVM) are of interest to real time adaptive array beam forming. Feasibility of high speed compact signal processing for applications such as sidelobe cancelling has yet to be demonstrated. The importance of the OMVM stems from the following:

(1) Because of its optical parallelism it can operate at extremely high speeds, surpassing by several orders of magnitude competing electronic techniques.

(2) The matrix-vector multiplication operation is basic to a variety of important signal processing applications, including adaptive array processing and other general control operations.

The very high speed of the OMVM strongly suggests its application to iterative processing, where in a very short time the system can cycle through perhaps 100 iterations, and assuming necessary convergence, yield a solution. This method has been suggested as a means of inverting the covariance matrix associated with noise samples from adaptive array antenna elements. A review of the jammer problem for radars indicates that solution is required in the 1 μsec to 10 μsec range. The spatial light modulator must be updated within this time frame.

We will describe here the optical architecture of an incoherent matrix vector multiplier where the matrix is complex and separable into outer product vectors.

INTRODUCTION

In the late 1970's, researchers at Stanford University began investigation of a new class of incoherent electro-optical processors for high speed matrix-vector multipliers (ref. 1, 2). Work has continued, both at Stanford (ref. 3) and Carnegie Mellon Institute, to the point were significant processing capabilities have been demonstrated. It is not unreasonable to expect that such systems could perform 10^8 matrix-vector multiplies per second with 100x100 element matrices.

One limitation of these processors as developed thus far is the relatively fixed nature of the mask representing the matrix. In general photographic masks have been employed, and processing where the matrix must change with some rapidity has been impossible.

In our search for a matrix mask that can be updated rapidly, we have considered several candidate spatial light modulators (ref. 4) (SLMs), including the Itek PROM, and the Hughes LCLV. Neither the PROM nor the LCLV can be cycled at rates approaching 1 MHz, and these candidates are thus unacceptable. A mosaic of pockels cells could perhaps be made to operate in the desired fashion. However, there are the usual problems of switching high voltages in microseconds, and severe electronic crosstalk appears inevitable. We have settled on a scheme whereby the mask is implemented acousto-optically.

Acousto-optic modulators possess several desirable features when used in signal processing systems. These include reliability, availability, high response speeds, and wide bandwidths. However, they are not easily adaptable to use as 2-D spatial light modulators. The basic technique and system architecture for a pair of 1-D SLMs is discussed in the following section.

ACOUSTO-OPTIC IMPLEMENTATION OF VECTOR MULTIPLICATION BY OUTER PRODUCT MATRICES

The operation to be implemented is:

$$\overline{Y} = \underline{H}\,\overline{X} \qquad (1)$$

where $\overline{X}$ is a complex input column vector.

$\underline{H}$ is a complex matrix separable into outer products:

$$\underline{H} = \overline{A}\,\overline{B}^t \qquad (2)$$

The multiplication of separable matrices would be possible if the OMVM were a coherent processor. When the necessary complex to non-negative real encoding of $\underline{H}$ is implemented for an incoherent processor, the separable form of the system matrix is destroyed. However by using time multiplexing we can still avoid the requirement for 2-D masks. Consider one possible method. We can define expressions in which vectors and matrices are defined in the following form (ref 5,6).

$$\overline{X} = \overline{X}_0 + \overline{X}_1 e^{i\theta} + \overline{X}_2 e^{i2\theta} \qquad (3a)$$

$$\overline{Y} = \overline{Y}_0 + \overline{Y}_1 e^{i\theta} = \overline{Y}_2 e^{i2\theta} \qquad (3b)$$

$$\underline{H} = \underline{H}_0 + \underline{H}_1 e^{i\theta} + \underline{H}_2 e^{i2\theta} \qquad (3c)$$

where $\overline{X}_0$, $\overline{X}_1$, $\overline{X}_2$; $\overline{Y}_0$, $\overline{Y}_1$, $\overline{Y}_2$; $\underline{H}_0$, $\underline{H}_1$, $\underline{H}_2$ are all non-negative real and where $\theta = 2\pi/3$. The components $\overline{X}_i$, $\overline{Y}_i$, $\overline{H}_i$, i = 0,1,2 have the dimensions of $\overline{X}$, $\overline{Y}$, and $\underline{H}$, respectively.

The basic matrix vector multiplication operation can be written:

$$\begin{bmatrix} \overline{Y}_0 \\ \overline{Y}_1 \\ \overline{Y2} \end{bmatrix} = \begin{bmatrix} \underline{H}_0 & \underline{H}_2 & \underline{H}_1 \\ \underline{H}_1 & \underline{H}_0 & \underline{H}_2 \\ \underline{H}_2 & \underline{H}_1 & \underline{H}_0 \end{bmatrix} \begin{bmatrix} \overline{X}_0 \\ \overline{X}_1 \\ \overline{X}_2 \end{bmatrix} \qquad (4)$$

Consider the case where $\underline{H}$ is an outer product matrix (2). If we express (2) in terms of a three component decomposition, we have $\underline{H} = (\overline{A}_0 + \overline{A}_1 e^{i\theta} + \overline{A}_2 e^{i2\theta})\ (\overline{B}_0^t + \overline{B}_1^t e^{i\theta} + \overline{B}_2^t e^{i2\theta})^t$ (5)

Multiplying through and collecting terms in powers of exp (iθ) we find that

$$\begin{aligned} \underline{H}_0 &= \overline{A}_0 \overline{B}_0 + \overline{A}_1 \overline{B}_2 + \overline{A}_2 \overline{B}_1 \\ \underline{H}_1 &= \overline{A}_0 \overline{B}_1 + \overline{A}_1 \overline{B}_0 + \overline{A}_2 \overline{B}_2 \qquad (6) \\ \underline{H}_2 &= \overline{A}_0 \overline{B}_2 + \overline{A}_1 \overline{B}_1 + \overline{A}_2 \overline{B}_0 \end{aligned}$$

To represent $\underline{H}$ by a three-component decomposition, we need nine outer products.

These nine outer products can be produced using a pair of crossed acousto-optic cells with multiple transducers. Figure 1 illustrates the case where $\overline{A}\,\overline{B}^t$ is a two-dimensional complex matrix. The desired outer products are formed at the intersection of the acoustic waves, in the following format: Let $\underline{M}$ =

$$\underline{M} = \begin{bmatrix} \overline{A}_0\,\overline{B}_0 & \overline{A}_0\,\overline{B}_1 & \overline{A}_0\,\overline{B}_2 \\ \overline{A}_1\,\overline{B}_0 & \overline{A}_1\,\overline{B}_1 & \overline{A}_1\,\overline{B}_2 \\ \overline{A}_2\,\overline{B}_0 & \overline{A}_2\,\overline{B}_1 & \overline{A}_2\,\overline{B}_2 \end{bmatrix} \qquad (7)$$

At this stage we encounter a difficulty in system architecture: although the various outer products of (3), together with input vector components $\overline{X}_0$, $\overline{X}_1$, $\overline{X}_2$, are sufficient to allow calculation of $\overline{Y}$, these outer products are not arranged so as to allow direct implementation of equation (4). The three-component form of matrix H

$$\underline{H} = \begin{bmatrix} \underline{H}_0 & \underline{H}_2 & \underline{H}_1 \\ \underline{H}_1 & \underline{H}_0 & \underline{H}_2 \\ \underline{H}_2 & \underline{H}_1 & \underline{H}_0 \end{bmatrix} \qquad (8)$$

is itself not representable as an outer product. As a consequence, it appears as though system architecture will of necessity be more complicated than would be the case were $\underline{H}$ directly realized by, for example, a more general 2-D SLM.

Let us consider the problem further to gain a better understanding of its nature. We simplify notation by denoting the elements of equation (7) by $\underline{M}_{ij}$:

$$\underline{M} = \begin{bmatrix} \underline{M}_{00} & \underline{M}_{01} & \underline{M}_{02} \\ \underline{M}_{10} & \underline{M}_{11} & \underline{M}_{12} \\ \underline{M}_{20} & \underline{M}_{21} & \underline{M}_{22} \end{bmatrix} \qquad (9)$$

From equations (3), (6), (7) and (9) we have

$$\bar{Y}_0 = \underline{H}_0\bar{X}_0 + \underline{H}_2X_1 + \underline{H}_1\bar{X}_2 \tag{10a}$$

$$= \underline{M}_{00}\bar{X}_0 + \underline{M}_{12}\bar{X}_0 + \underline{M}_{21}\bar{X}_0 +$$

$$\underline{M}_{02}\bar{X}_1 + \underline{M}_{11}\bar{X}_1 + \underline{M}_{20}\bar{X}_1 +$$

$$\underline{M}_{01}\bar{X}_2 + \underline{M}_{10}\bar{X}_2 + \underline{M}_{22}\bar{X}_2$$

$$\bar{Y}_1 = \underline{H}_1\bar{X}_0 + \underline{H}_0\bar{X}_1 + \underline{H}_2\bar{X}_2 \tag{10b}$$

$$\underline{M}_{01}\bar{X}_0 + \underline{M}_{10}\bar{X}_0 + \underline{M}_{22}\bar{X}_0 +$$

$$\underline{M}_{00}\bar{X}_1 + \underline{M}_{12}\bar{X}_1 + \underline{M}_{21}\bar{X}_1 +$$

$$\underline{M}_{02}\bar{X}_2 + \underline{M}_{11}\bar{X}_2 + \underline{M}_{20}\bar{X}_2$$

$$\bar{Y}_2 = \underline{H}_2\bar{X}_0 + \underline{H}_1\bar{X}_1 + \underline{H}_0\bar{X}_2 \tag{10c}$$

$$\underline{M}_{02}\bar{X}_0 + \underline{M}_{11}\bar{X}_0 + \underline{M}_{20}\bar{X}_0 +$$

$$\underline{M}_{01}\bar{X}_1 + \underline{M}_{10}\bar{X}_1 + \underline{M}_{22}\bar{X}_1 +$$

$$\underline{M}_{00}\bar{X}_2 + \underline{M}_{12}\bar{X}_2 + \underline{M}_{21}\bar{X}_2$$

Inspection shows that component $\bar{Y}_n$ can be expressed by

$$\bar{Y}_n = \sum_{i,j,k} \underline{M}_{ij}\bar{X}_k \tag{11}$$

where i,j,k satisfy the relationship

$$(i+j+k) \bmod 3 = n$$

Given the matrix $\underline{M}$ and input $\bar{X} = (\bar{X}_0\ \bar{X}_1\ \bar{X}_2)^t$, it is not possible to generate all the terms of (10) by a single vector-matrix multiplication operation. It is, however, possible to generate $\bar{Y}$ with three vector-matrix multiplication operations, as follows.

First, input the vector $(\bar{X}_0\ \bar{X}_2\ \bar{X}_1)^t$ and calculate its product with $\underline{M}$. We write the resulting product as

$$\begin{matrix} \bar{Y}_{0a} \\ \bar{Y}_{1a} \\ \bar{Y}_{2a} \end{matrix} = \underline{M} \begin{matrix} \bar{X}_0 \\ \bar{X}_2 \\ \bar{X}_1 \end{matrix} \qquad (12a)$$

This is followed by the operations

$$\begin{matrix} \bar{Y}_{1b} \\ \bar{Y}_{2b} \\ \bar{Y}_{0b} \end{matrix} = \underline{M} \begin{matrix} \bar{X}_1 \\ \bar{X}_0 \\ \bar{X}_2 \end{matrix} \qquad (12b)$$

$$\begin{matrix} \bar{Y}_{2c} \\ \bar{Y}_{0c} \\ \bar{Y}_{1c} \end{matrix} = \underline{M} \begin{matrix} \bar{X}_2 \\ \bar{X}_1 \\ \bar{X}_0 \end{matrix} \qquad (12c)$$

If the outputs of these three operations are properly combined, the result is the desired three-component representation for Y. Specifically,

$$\begin{aligned} \bar{Y}_0 &= \bar{Y}_{0a} + \bar{Y}_{0b} + \bar{Y}_{0c} \qquad (13) \\ \bar{Y}_1 &= \bar{Y}_{1a} + \bar{Y}_{1b} + \bar{Y}_{1c} \\ \bar{Y}_2 &= \bar{Y}_{2a} + \bar{Y}_{2b} + \bar{Y}_{2c} \end{aligned}$$

Study of these equations shows that as the components $\overline{X}_0$, $\overline{X}_2$, and $\overline{X}_1$ are stepped cyclically in the input, the components of $\overline{Y}_0$, $\overline{Y}_1$, and $\overline{Y}_2$ are themselves stepped cyclically. One method for performing the summation of (12) and (13) is to use charge-coupled integrating detectors in the output plane of the optical system. The integrated charges can then be stepped synchronously with the input components. Another method is to use high speed analog switches and short delay lines for storage.

PLANS FOR IMPLEMENTATION

The basic optical layout of the iterative processor is shown schematically in figure 2. Light emitted from the sources is collimated into N parallel beams by means of a set lens L1. Each lens is a segment of a spherical lens. Thus, each row of the first SLM is illuminated by a collimated beam from one of the light sources. The output beams from the AO Bragg cell, A, are focused by lens L4, spatially filtered by S1, and recollimated by L5. The lens set L4-L5 images the aperture of Bragg cell A at the second Bragg cell, B. The modulated beam from the second Bragg cell is spatially filtered and brought to focus on the linear detector array. The spatial fitlers are necessary only to remove the zero and second order beams. The output beams are modulated in intensity at NxN points within the optical aperture. This representation is equivalent to having both masks located at the same plane. The outputs from the light emitters are modulated by vectors whose elements are $\overline{X}_j$. On multiplication, the output terms are $\overline{Y}_i = \sum_j \overline{A}_j \, \overline{B}_i \, \overline{X}_j$.

When a non-negative real approach is implemented, the above scheme must be modified. The speed of the processor is slowed down by approximately a factor of three, since for each iteration of three successive multiplications must be carried out. After each multiplication the analog data must be stored while the input vector is shifted. The three successive outputs are added as in equation (18). The output vector then becomes the input to the light emitters as in other iterative optical processors.

The time multiplexed method described can be implemented using CCD shift registers. Unfortunately, state of the art CCD shift registers operating at the necessary speed lack the dynamic range and noise properties necessary for our applications. We have thus chosen to implement the previously mentioned procedure utilizing fast four channel CMOS analog multiplexer switches arranged in a switching matrix configuration. Referring to figure 3, as the MUX is clocked, its output is shifted along the linear laser diode arrays. Similarly, at the linear photodiode array, the outputs $(\overline{Y}_{0a}\ \overline{Y}_{1a}\ \overline{Y}_{2a})^t$ from the first input cycle $(\overline{X}_0\ \overline{X}_2\ \overline{X}_1)^t$ are stored in coaxial cable delay lines. The input X vector is then switched to $(\overline{X}_1\ \overline{X}_0\ \overline{X}_2)^t$ and the output $(\overline{Y}_{1b}\ \overline{Y}_{2b}\ \overline{Y}_{0b})^t$ is switched to the proper location and stored in another delay line. This procedure is repeated a third time, with the appropriate combinations taken to yield the desired outputs.

Some of the problems particular to the implementation involve non-linearities in the Bragg diffraction process, electrical and acoustic crosstalk, and limited dynamic range due to light scattering.

CONCLUSIONS

An analysis has been performed of the applicability of iterative optical processors in environments requiring a short transient response time. An approach has been proposed utilizing realizable acousto-optic single dimension light modulators where the matrix is the outer product of two vectors. The problems associated with a non-negative real representation have been addressed. A solution has been proposed albeit at some cost of processor speed. A proof of concept experiment is underway which will better determine the limitations and advantages of this method.

REFERENCES

1. Goodman, J. et al: SPIE, Vol. 190, 1979, p. 484.
2. Psaltis D., Casasant, D.; and Carlotto, M.: Iterative Optical Processor (IOP) for Adaptive Phase Array Radar Processing. SPIE, Vol. 180, 1979, p. 114.
3. Dias, A. R.: Incoherent Optical Matrix-Vector Multiplication for High Speed Data Processing. Ph.D. Thesis, Stanford University, 1980.
4. Casasant, D.: Spatial Light Modulators and Their Use in Optical Data Processing. SPIE, Vol. 128, 1977, p. 56.
5. Goodman, J. W.; and Woody, L. M.: Method for Performing Complex Valued Linear Operations on Complex Valued Data Using Incoherent Light. Applied Optics, Vol. 16, No. 10, Oct. 1977, p. 2611.
6. Ref. 3, p. 13.

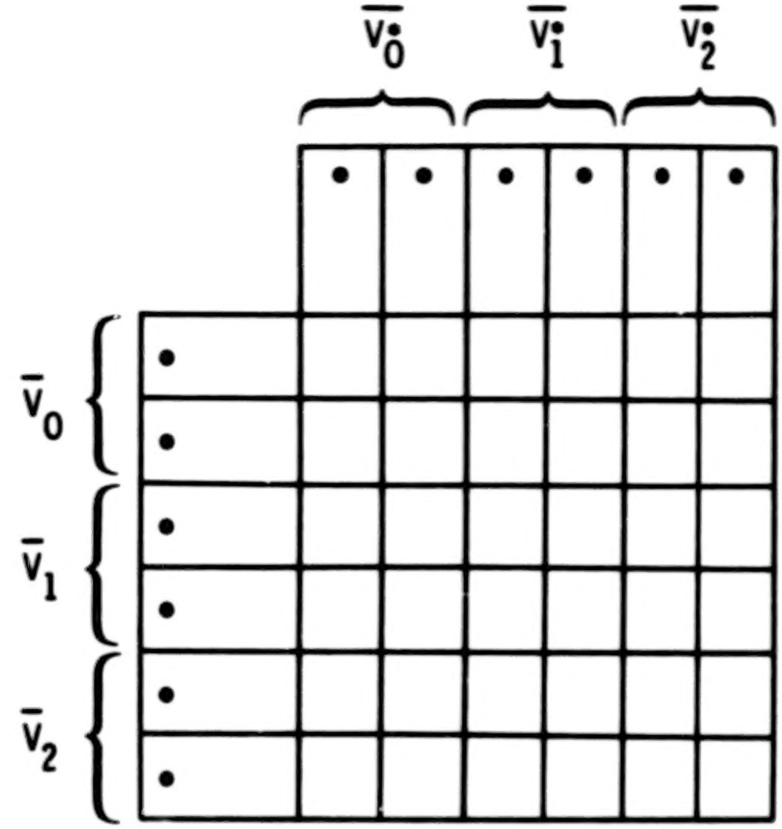

$\bar{V}_0^\bullet$
$\bar{V}_1^\bullet$
$\bar{V}_2^\bullet$
$\bar{V}_0$
$\bar{V}_1$
$\bar{V}_2$

Figure 1.- Outer product geometry for producing elements of $\underline{M}$ with a pair of six transducer acousto-optic cells.

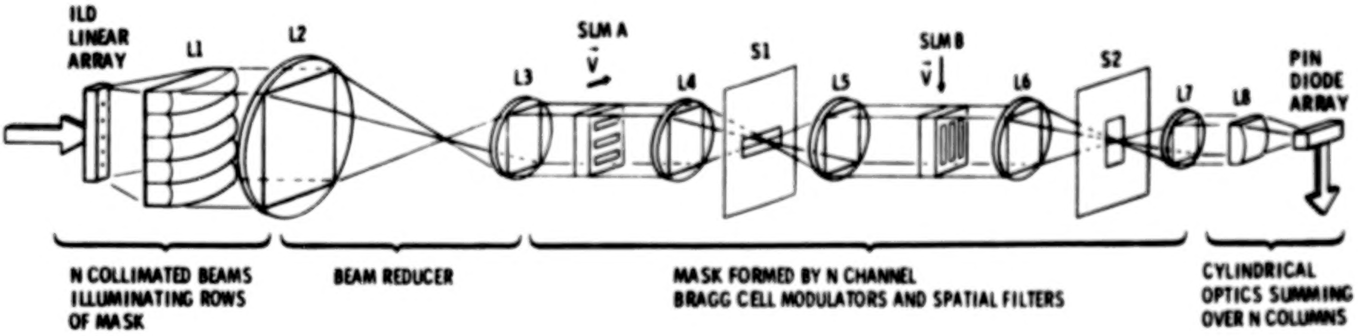

ILD
LINEAR
ARRAY
L1
L2
L3
SLM A
$\bar{V}$
L4
S1
L5
SLM B
$\bar{V}$
L6
S2
L7
L8
PIN
DIODE
ARRAY
N COLLIMATED BEAMS ILLUMINATING ROWS OF MASK
BEAM REDUCER
MASK FORMED BY N CHANNEL BRAGG CELL MODULATORS AND SPATIAL FILTERS
CYLINDRICAL OPTICS SUMMING OVER N COLUMNS

Figure 2.- Optical system for matrix vector processor with separable masks.

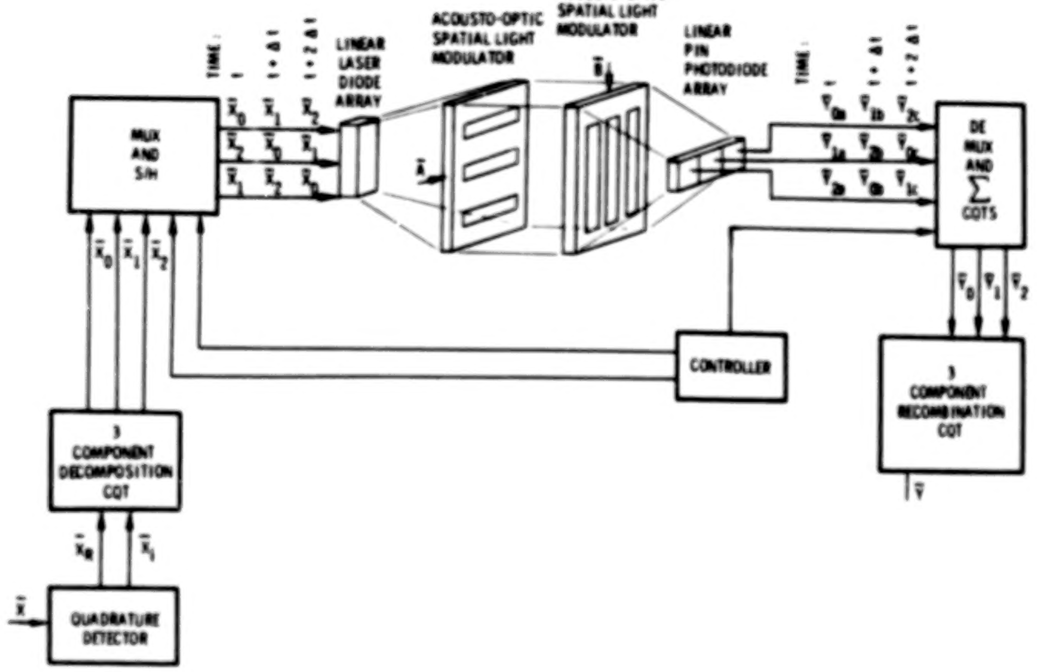

TIME: t t + Δt t + 2Δt
LINEAR LASER DIODE ARRAY
ACOUSTO-OPTIC SPATIAL LIGHT MODULATOR
ACOUSTO-OPTIC SPATIAL LIGHT MODULATOR
LINEAR PIN PHOTODIODE ARRAY
TIME: t t + Δt t + 2Δt
MUX AND S/H
$\bar{x}_0$ $\bar{x}_1$ $\bar{x}_2$
$\bar{x}_2$ $\bar{x}_0$ $\bar{x}_1$
$\bar{x}_1$ $\bar{x}_2$ $\bar{x}_0$
$\bar{A}$
$\bar{B}$
$\bar{y}_{0a}$ $\bar{y}_{1b}$ $\bar{y}_{2c}$
$\bar{y}_{1a}$ $\bar{y}_{2b}$ $\bar{y}_{0c}$
$\bar{y}_{2a}$ $\bar{y}_{0b}$ $\bar{y}_{1c}$
DE MUX AND Σ COTS
$\bar{y}_0$ $\bar{y}_1$ $\bar{y}_2$
$\bar{x}_0$ $\bar{x}_1$ $\bar{x}_2$
CONTROLLER
3 COMPONENT RECOMBINATION CQT
$\bar{y}$
3 COMPONENT DECOMPOSITION CQT
$\bar{x}_R$ $\bar{x}_I$
$\bar{x}$
QUADRATURE DETECTOR

Figure 3.- System architecture for OMVM, with time varying mask.

INCOHERENT OPTICAL MATRIX-MATRIX MULTIPLIER

A.R. Dias
Radar and Optics Division
Environmental Research Institute of Michigan
P.O. Box 8618, Ann Arbor, Michigan 48017

INTRODUCTION

In recent years a growing interest has developed in incoherent optical processing (ref. 1). As the invention of the laser was an important force in coherent optics, the advent of fast and very compact solid state light sources is a major force behind this interest.

Incoherent optical processing has indeed several appealing characteristics, namely higher immunity to optical noise and reduced dynamic range requirements than coherent processing as well as an inherent versatility of input format (serial or parallel).

A very fast fully parallel incoherent optical multiplier was developed at Stanford (refs. 2 and 3) and this device has been shown to be capable of performing complex-valued arithmetic on vector formatted data arriving in parallel to the processor inputs.

There is a considerable number of applications for processors of this sort however there is a need to extend their capability for two-dimensional parallel data. Indeed one can understand the interest in speeding real time processing capabilities in the context of reconnaissance sensors, two-dimensional beamforming, etc.

This work addresses the latter case. We propose an extension of the concepts developed earlier to the new generation of matrix-matrix multipliers using incoherent light. This new technology has the potential of addressing problems involving two-dimensional mathematical transforms, two-dimensional pattern recognition, and high-speed processing of synthetic aperture radar data among others.

SYSTEM ARCHITECTURE

Let us then consider the matrix multiplication

$$\underline{C} = \underline{A}\,\underline{B} \tag{1}$$

where $\underline{A}$ is an M x K matrix, and $\underline{B}$ and $\underline{C}$ are K x N and M x N matrices respectively.

The element c_{mn} in the m-th row and n-th column is given in terms of the elements a_{mk} of $\underline{A}$ and b_{kn} of $\underline{B}$ by

$$c_{mn} = \sum_{k=0}^{K-1} a_{mk}\, b_{kn} \tag{2}$$

A matrix can, of course, be viewed as an ordered collection of column vectors.

If we define

$$\underline{B} = \sum_{n=0}^{N-1} \underline{B_n} \text{ where } \underline{B_n} = \begin{bmatrix} 0 & 0 & \cdots & b_{0n} & \cdots & 0 \\ 0 & 0 & \cdots & b_{1n} & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ 0 & \cdots & & b_{K-1,n} & \cdots & 0 \end{bmatrix} \tag{3a}$$

and

$$\underline{C} = \sum_{n=0}^{N-1} \underline{C_n} \text{ where } \underline{C_n} = \begin{bmatrix} 0 & 0 & \cdots & c_{0n} & \cdots & 0 \\ 0 & 0 & \cdots & c_{1n} & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ 0 & \cdots & & c_{M-1,n} & \cdots & 0 \end{bmatrix} \tag{3b}$$

the matrix product $\underline{C} = \underline{A}\,\underline{B}$ becomes the sum of N matrices $\underline{C_n}$ whose only non-zero elements are in its n-th column. In consequence, the matrix product $\underline{C}$ can be seen as an ordered collection of column vectors each of which is the product of $\underline{A}$ by the corresponding column vector of $\underline{B}$. A matrix-matrix multiplication can then be expressed as a series of matrix-vector multiplications. The technology developed for the incoherent optical matrix-vector multiplier (ref. 2) can then be extended to this new application.

The concept here is to combine N matrix-vector multipliers in a single processor. The physical realization of this architecture is depicted in Figure 1.

The elements of matrix $\underline{B}$* are entered as irradiance values of a two-dimensional incoherent array that may be realized using narrow-band light-emitting diodes (LED), laser diodes, or even an array of polished optical fiber ends. The elements of matrix $\underline{A}$ are encoded as transmission values of an optical transparency or as reflected values of an optically reflective device. In either case, its implementation can be fixed in time (e.g., a film transparency) or it can vary with time (e.g., a real time optical modulator).

The first block of optics (figure 1) images horizontally the input array $\underline{B}$ onto the optical mask $\underline{A}$ and spreads vertically the light of each light source along the corresponding vertical column. Figure 2 illustrates the illumination produced by the central column of a 3 x 3 input matrix on a 3 x 3 mask. The light transmitted through the mask encodes all the possible inner products of the column vectors of the input matrix by all the row vectors of the stored matrix.

The second block of optics (figure 1) integrates these inner products and focuses the light on the proper output array locations -- imaging vertically and focusing horizontally onto the vertical column corresponding to the particular horizontal row at the input. This is a demanding operation that can be more clearly visualized in the case of a 3 x 3 matrix. We illustrate in figure 3 how we obtain the outputs (1, 1) and (3, 2). The output (3, 2) can readily be seen as the inner product of the 3-rd row of matrix A and the 2-nd column of matrix B as would be expected. Finally, figure 4 shows the light path through a particular cell in the stored mask. The element a_{mk} (the central element in this example) maps the elements of the k-th row of input B onto the elements of the m-th row of output C.

---

*At this stage, we assume that the elements of $\underline{A}$ and $\underline{B}$ are real and non-negative. Later, we will remove this restriction.

OPTICAL IMPLEMENTATION

The optics block labeled "Optics I" in figure 1 has the same requirements as the mask illumination of the incoherent optical matrix-vector multiplier (ref. 3). In the latter case, a discrete set of classical discrete optical elements in one version and an array of multimode planar waveguides in another version were used. In the present case, a very similar form of the classical discrete optical components approach mentioned could certainly be used.

The second block of optics -- Optics II -- has more demanding requirements. In effect, it must steer the light transmitted by each cell in the stored mask to all N elements in the corresponding row of the output detector array (figure 4). Holographic optical elements can be used as the steering beam devices in this type of processor. An example of the application of this technology was shown by S. Case and his co-workers in the paper "Multifacet Holographic Optical Elements," presented at the Optical Society of America Annual Meeting in October 1980 (Abstract in J.O.S.A., 12, 1980).

In this context, we suggest the geometry depicted in figure 5 where each narrowband light source and each cell of the stored mask are located behind a holographic optical element (HOE). This is a well established technology with the potential for a highly efficient light management.

COMPLEX-VALUED ARITHMETIC

It is, of course, important for a processor to have the capability to perform complex-valued operations. The incoherent optical processor we report here manipulates light intensities which are real-valued and non-negative in nature. Methods for encoding complex numbers similar to those used in the incoherent optical matrix-vector multiplier (ref. 4) can be applied here.

The complex-valued arithmetic capability built in this processor is an immediate extension of the schemes utilized before, and for completeness, we will briefly elaborate two possible methods.

A. Three Vector Decomposition

Each complex quantity is decomposed uniquely along the three phasors $1 \cdot \exp[j0]$, $1 \cdot \exp[j2\pi/3]$ and $1 \cdot \exp[j4\pi/3]$. There is more than one way to perform this decomposition (ref. 4, Appendix A) but we do not need to dwell on it here.

The three matrices can be decomposed according to

$$\underline{A} = \underline{A_0}\, e^{j0} + \underline{A_1}\, e^{j\frac{2\pi}{3}} + \underline{A_2}\, e^{j\frac{4\pi}{3}}$$

$$\underline{B} = \underline{B_0}\, e^{j0} + \underline{B_1}\, e^{j\frac{2\pi}{3}} + \underline{B_2}\, e^{j\frac{4\pi}{3}} \tag{4}$$

$$\underline{C} = \underline{C_0}\, e^{j0} + \underline{C_1}\, e^{j\frac{2\pi}{3}} + \underline{C_2}\, e^{j\frac{4\pi}{3}}$$

where the elements of the different matrix components are all real-valued non-negative quantities.

The matrix-matrix multiplication $\underline{C} = \underline{A}\,\underline{B}$ can now be expressed in terms of these quantities as

$$\begin{bmatrix} \underline{C_0} \\ \underline{C_1} \\ \underline{C_2} \end{bmatrix} = \begin{bmatrix} \underline{A_0} & \underline{A_2} & \underline{A_1} \\ \underline{A_1} & \underline{A_0} & \underline{A_2} \\ \underline{A_2} & \underline{A_1} & \underline{A_0} \end{bmatrix} \begin{bmatrix} \underline{B_0} \\ \underline{B_1} \\ \underline{B_2} \end{bmatrix} \tag{5}$$

from which the output complex-valued matrix $\underline{C}$ can be obtained after the proper recombination of its real valued and non-negative components $\underline{C_0}$, $\underline{C_1}$ and $\underline{C_2}$.

The processor's configuration is shown in figure 6, where we have omitted the optics. The size of the input and output arrays increases by a factor of 3 and the stored mask by a factor of 9.

B. Biased Real and Imaginary Parts Decomposition

Another encoding scheme decomposes each complex quantity into its real and imaginary parts and adds to these a bias term.

$\underline{A}$ and $\underline{B}$ can then be decomposed into their real and imaginary parts.

$$\underline{A} = \underline{A_R} + j\underline{A_I}$$

$$\underline{B} = \underline{B_R} + j\underline{B_I} \tag{6}$$

and we can express $\underline{C}$ as

$$\begin{bmatrix} \underline{C_R} \\ \underline{C_I} \end{bmatrix} = \begin{bmatrix} \underline{A_R} & -\underline{A_I} \\ \underline{A_I} & \underline{A_R} \end{bmatrix} \begin{bmatrix} \underline{B_R} \\ \underline{B_I} \end{bmatrix} \tag{7}$$

Assuming known the dynamic range of the elements of the real and imaginary parts of $\underline{A}$ and $\underline{B}$, we assign a bias matrix $\underline{B_A}$ to matrix $\underline{A}$ and a bias matrix $\underline{B_B}$ to matrix $\underline{B}$.

These bias matrices have constant elements. All elements in the same column are identical, but they can vary from column to column.

If now we configure the system as it is shown in figure 7, the detected outputs $\underline{C_R}$ and $\underline{C_I}$ are given by

$$\underline{C'_R} = \underline{A_R B_R} - \underline{A_I B_I} + \left[\underline{A_R B_B} - \underline{A_I B_B}\right] + \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{8a}$$

$$\underline{C'_I} = \underline{A_I B_R} + \underline{A_R B_I} + \left[\underline{A_I B_B} + \underline{A_R B_B}\right] + \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{8b}$$

The real and imaginary parts $\underline{C_R}$ and $\underline{C_I}$ of the product $\underline{C}$ (given by equation (4)) can be obtained from equations (8a) and (8b)

$$\underline{C_R} = \underline{C'_R} - \left[\underline{A_R B_B} - \underline{A_I B_B}\right] - \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{9a}$$

$$\underline{C_I} = \underline{C'_I} - \left[\underline{A_I B_B} + \underline{A_R B_B}\right] - \underline{B_A}\left[\underline{B_R} + \underline{B_B} + \underline{B_I} + \underline{B_B}\right] \tag{9b}$$

In each equation, the first bracket is known a priori and it can be electronically represented by a fixed voltage. The second bracket can be measured, adding an extra transparent row to the stored mask. Hence, with proper and straightforward post-detection electronics, one obtains the desired product real and imaginary parts.

SYSTEM PERFORMANCE ESTIMATE

At this stage of the system development, we address ourselves only to the major performance parameter -- the data throughput.

We can reasonably postulate that it will be feasible to build matrix-matrix multipliers with complex-valued matrices with sizes larger than 100 x 100. Without being overly optimistic, we can envision the electronics side of such processors being driven at clock rates greater than 100 MHz. This implies data throughput rates exceeding 10^{12} complex-valued samples per second. This is at least two degrees of magnitude higher than throughput rates obtained with matrix-vector multipliers.

APPLICATIONS

The parallel processing character of this processor is an essential feature that is extremely useful when a large number of input channels are present simultaneously.

A large number of problems dealing with matrix multiplication can be implemented with this processor. We may refer problems involving Fourier transforms (beam nulling, beamforming), image processing problems (pattern recognition, filtering), and signal identification (signature reconnaissance), to name just a few.

In this work, we will single out two fundamental mathematical applications that may be implemented using this technique.

A. Two-Dimensional Transforms

Let us consider a two-dimensional array $i(k, l)$ assuming values in the rectangular range $k = 0, 1, \ldots, N - 1$, $l = 0.1, \ldots, N - 1$. A two-dimensional transform over this array produces a second array $o(m, n)$ given by

$$o(m, n) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} t(m, n, k, l)\ i(k, l) \tag{10}$$

$$m = 0, 1, \ldots, N - 1$$

$$n = 0, 1, \ldots, M - 1$$

where $t(m, n, k, l)$ is the transform kernel.

We assume the kernel $t(m, n, k, l)$ is separable

$$t(m, n, k, l) = v(m, k)\ h(n, l). \tag{11}$$

We can then express $o(m, n)$ as

$$o(m, n) = \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} v(m, k)\ h(n, l)\ i(k, l) \tag{12}$$

and factorizing $h(n, l)$ and performing first the sum over k, we obtain

$$o(m, n) = \sum_{l=0}^{N-1} h(n, l) \sum_{k=0}^{N-1} v(m, k) i(k, l) \tag{13}$$

The sum over k is a matrix product (see equation (2)) producing the intermediate result

$$s(m, l) = \sum_{k=0}^{N-1} v(m, k) i(k, l) \tag{14}$$

which can be written in matrix notation

$$\underline{S} = \underline{V}\,\underline{I} \tag{15}$$

where $\underline{I}$ is the input matrix, $\underline{V}$ is the transform kernel matrix along the columns and $\underline{S}$ is the intermediate matrix. The output can then be expressed as

$$o(m, n) = \sum_{l=0}^{M-1} h(n, l)s(m, l) \tag{16}$$

or in matrix notation

$$\underline{O} = \left[\underline{H}\,(\underline{V}\,\underline{I})^T\right]^T \tag{17}$$

where $\underline{H}$ is the transform kernel matrix along the rows and $\underline{O}$ is the output matrix.

Equation (17) shows that two matrix multiplications in sequence can implement a two-dimensional transform (as given by equation (10)) when the transform kernel is separable.

The immediate application that comes to mind is the two-dimensional discrete Fourier transform (DFT). The discrete Fourier transform kernel is given by

$$t(m, n, k, l) = \exp\ -j\,\frac{2\pi mk}{N}\ \exp\ -j\,\frac{2\pi nl}{M} \tag{18}$$

$$v(m, k) = \exp\ -j\,\frac{2\pi mk}{N} \tag{19a}$$

$$h(n, l) = \exp\ -j\,\frac{2\pi nl}{M} \tag{19b}$$

Figure 8 illustrates a basic configuration for a two-dimensional DFT implemented with two matrix-matrix multipliers.

B. Iterative Processing

Another significant application of incoherent optical matrix-matrix multipliers is in the area of iterative processing. Figure 9 shows a basic configuration.

There are several possible uses for iterative techniques. Important examples are the phase retrieval problem related to imaging through turbulence (ref. 5) and the implementation of some numerical methods for matrix inversion.

This work was performed under an ERIM internal research program.

REFERENCES

1. Monahan, M.A., K. Bromley and R.P. Bocker, "Incoherent Optical Correlators, Proc. IEEE, 65, 121 (1977).

2. Goodman, J.W., A.R. Dias and L.M. Woody, "Fully Parallel, High Speed Incoherent Optical Method for Performing Discrete Fourier Transforms," Optics Letters, 1 (1978).

3. Goodman, J.W., A.R. Dias, L.M. Woody and J. Erickson, "Application of Optical Communications Technology to Optical Information Processing," Proc. SPIE 190, LASL Conf. on Optics 79 (1979).

4. Dias, Antonio R., "Incoherent Matrix-Vector Multiplication for High Speed Data Processing" Stanford University Ph.D. Thesis, June 1980, University Microfilms, 80-24641, Ann Arbor, MI. Also published as Stanford Electronics Laboratories, Tech. Report No. L722-4 (June 1980).

5. Fienup, J.R., "Space Object Imaging Through the Turbulent Atmosphere," Optical Eng., 18-5, 529, 1979.

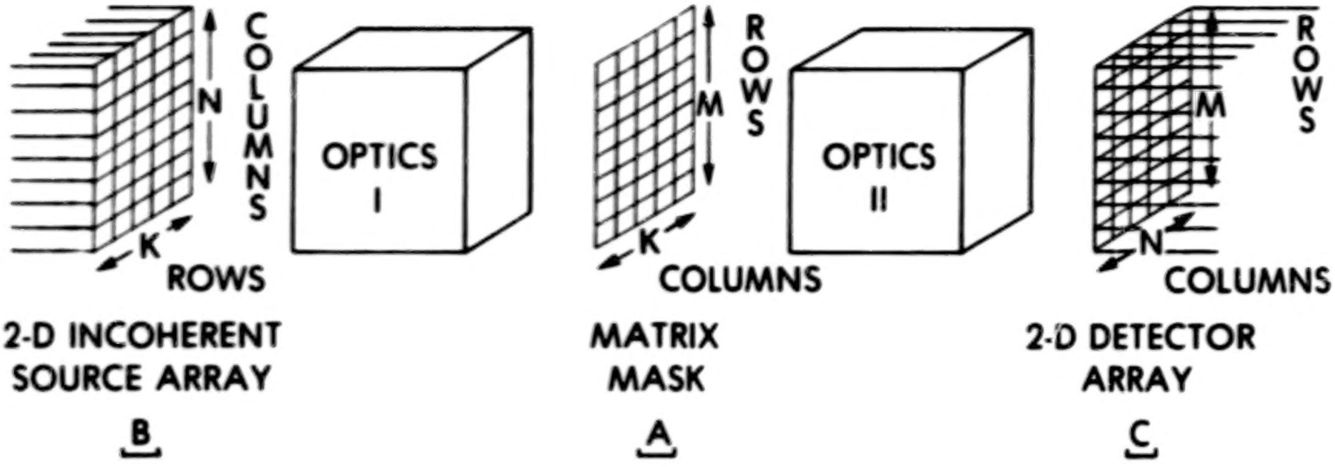

COLUMNS
N
K
ROWS
2-D INCOHERENT SOURCE ARRAY
B
OPTICS I
ROWS
M
K
COLUMNS
MATRIX MASK
A
OPTICS II
ROWS
M
N
COLUMNS
2-D DETECTOR ARRAY
C

Figure 1.- Basic block form of the incoherent optical matrix-matrix multiplier.

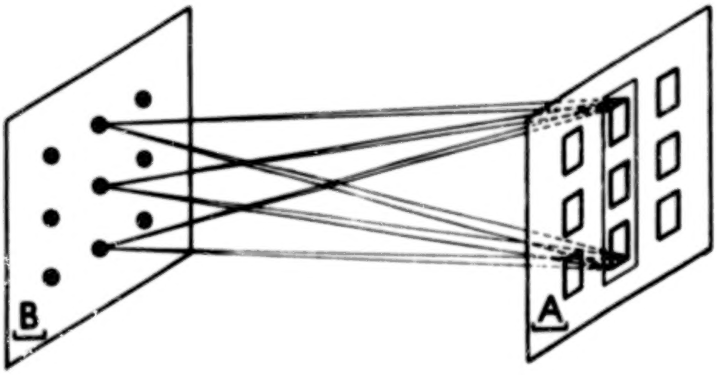


Figure 2.- Example of $\underset{\sim}{B}$ to $\underset{\sim}{A}$ illumination. For simplicity sake we will refer to matrix $\underset{\sim}{A}$ as the stored mask, though it may be variable with time.

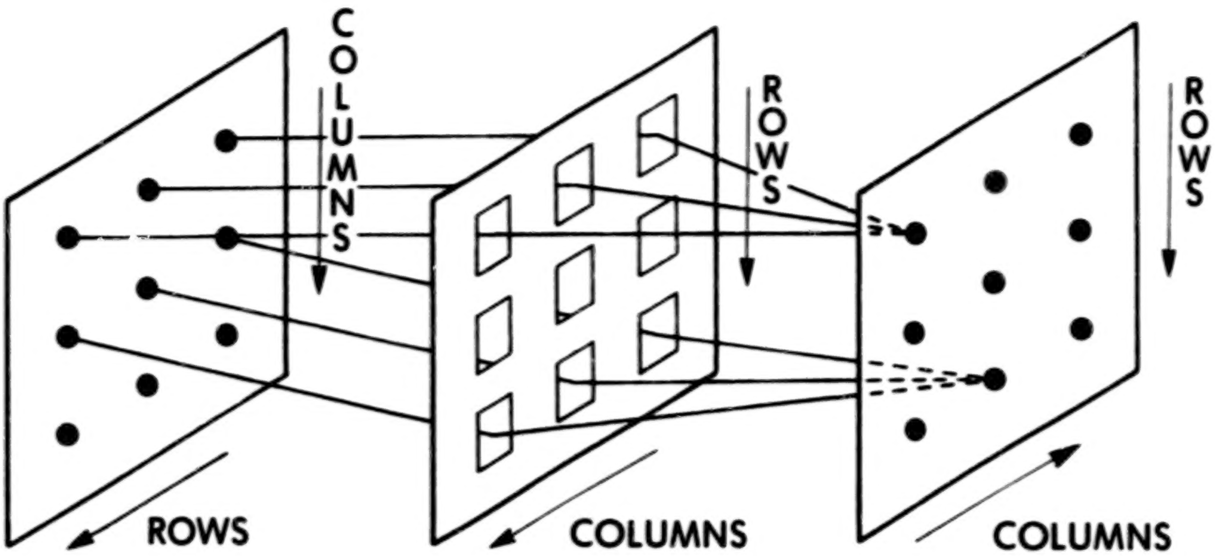

COLUMNS
ROWS
ROWS
ROWS
COLUMNS
COLUMNS

Figure 3.- Example of A to C illumination.

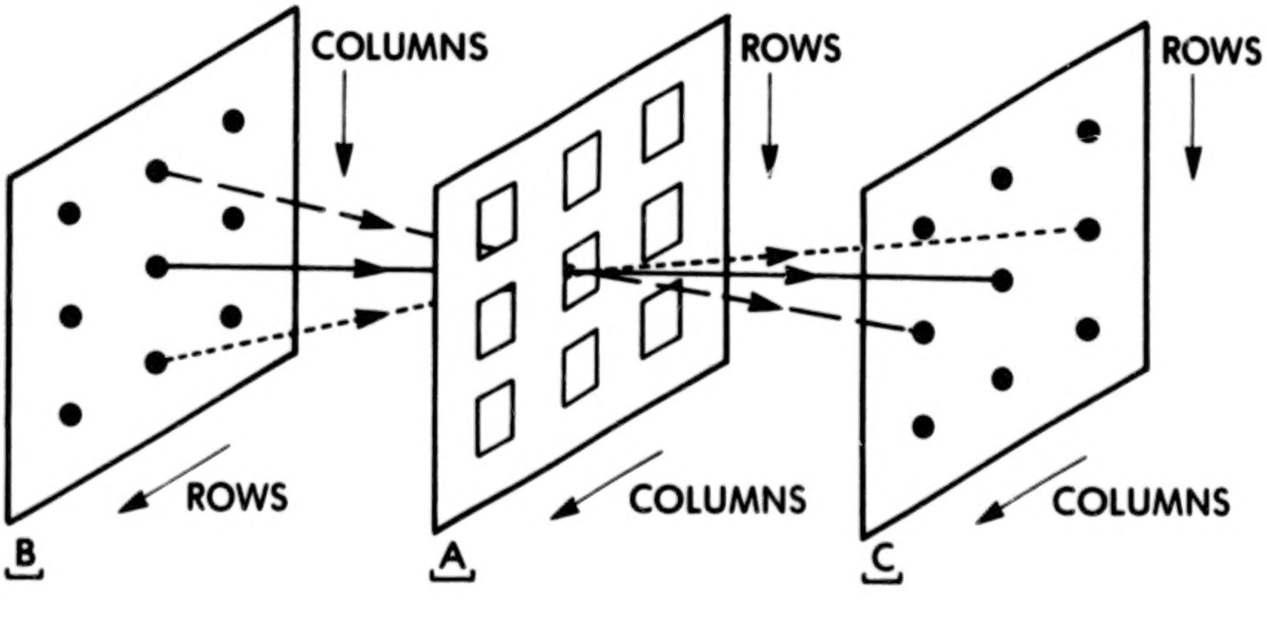

COLUMNS
ROWS
ROWS
ROWS
COLUMNS
COLUMNS
B
A
C

Figure 4.- Example of light path through an individual cell in the stored mask.

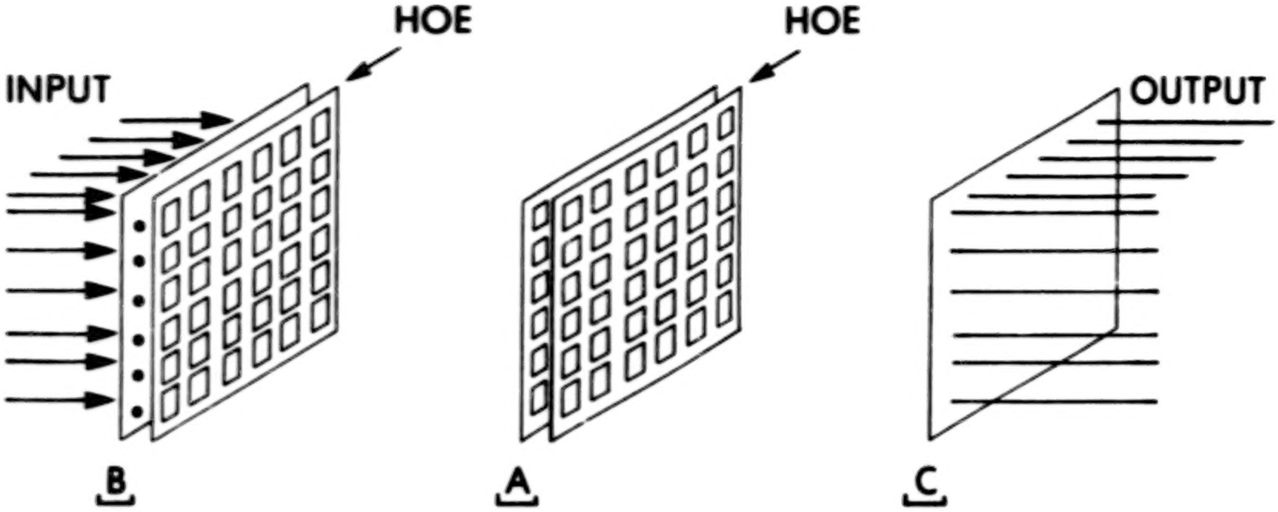

HOE
HOE
INPUT
OUTPUT
B
A
C

Figure 5.- Matrix-matrix multiplier. Implementation with holographic optical elements.

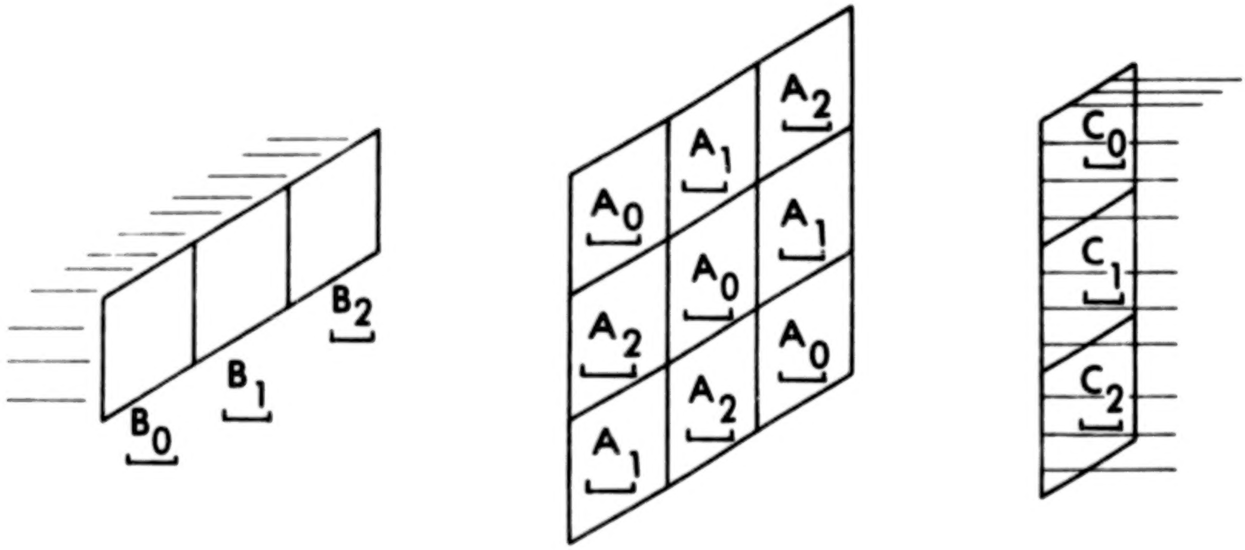
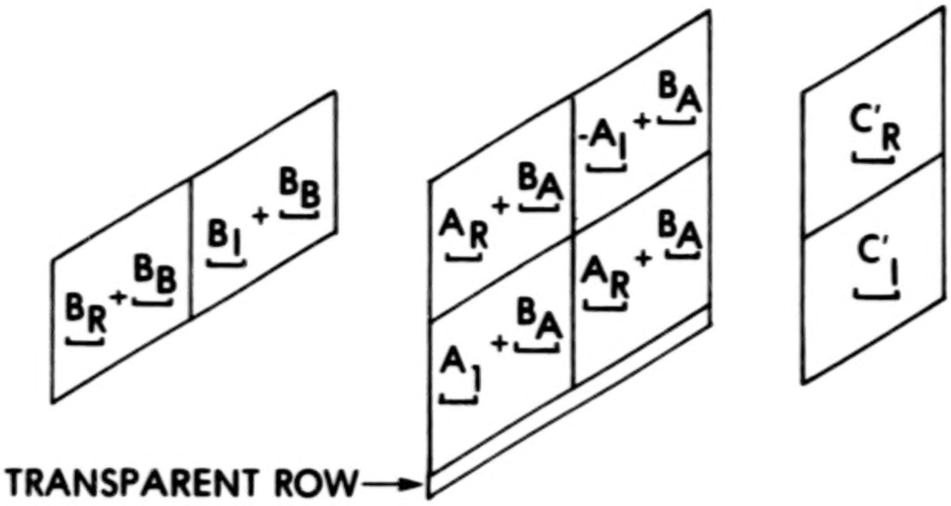

B0
B1
B2
A0
A1
A2
A1
A0
A2
A2
A1
A0
C0
C1
C2

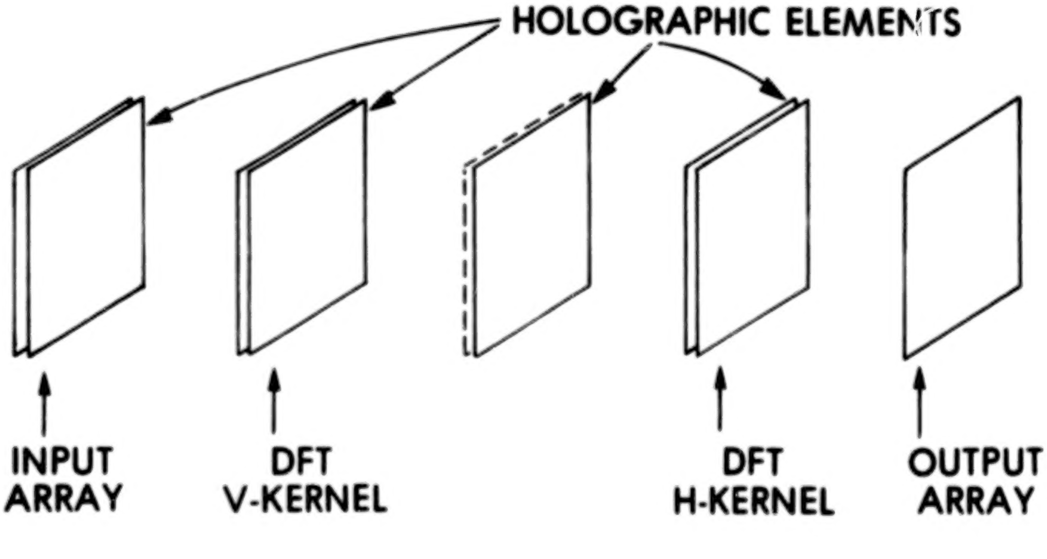
Figure 6.- Three-vector decomposition configuration.

BEST DOCUMENT AVAILABLE



$B_R + B_B$
$B_I + B_B$
$A_R + B_A$
$-A_I + B_A$
$A_R + B_A$
$A_I + B_A$
C'_R
C'_I
TRANSPARENT ROW

Figure 7.- Biased real and imaginary parts decomposition configuration.



HOLOGRAPHIC ELEMENTS
INPUT ARRAY
DFT V-KERNEL
DFT H-KERNEL
OUTPUT ARRAY

Figure 8.- 2D-FFT.

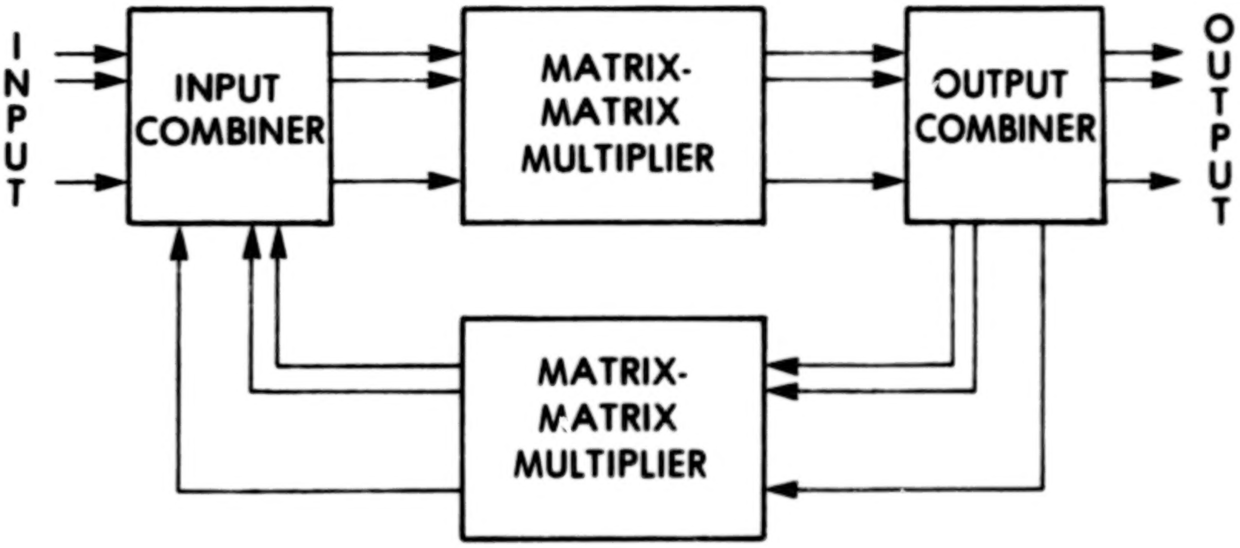

INPUT
INPUT COMBINER
MATRIX-MATRIX MULTIPLIER
OUTPUT COMBINER
OUTPUT
MATRIX-MATRIX MULTIPLIER

Figure 9.- Iterative processor.

Bulk and Integrated Acousto-Optic Spectrometers for Radio Astronomy

G. Chin, D. Buhl
Laboratory for Extraterrestrial Physics
Infrared and Radio Astronomy Branch

NASA/Goddard Space Flight Center
Greenbelt, MD 20771

J.M. Florez
Instrument Data Management Branch
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

One of the most significant developments in radio astronomy has been the recent discovery of over 50 different molecules in the interstellar medium. These observations have changed our picture of the distribution of mass in the galaxy, altered our understanding of the process of star formation, and also opened up a new and lively field of interstellar chemistry. This achievement was made possible not only by the development of sensitive heterodyne receivers (front-end) in the centimeter and millimeter range, but also by the construction of sensitive RF spectrometers (back-end) which enabled the spectral lines of molecules to be detected and identified. Traditionally, spectrometers have been constructed as banks of discrete adjacently tuned RF filters or as digital auto-correlators. However, a new technique combining acoustic bending of a collimated coherent light beam by a Bragg cell followed by detection by a sensitive array of photodetectors (thus forming an RF acousto-optic spectrometer (AOS)) promises to have distinct advantages over older spectrometer technology. An AOS has wide bandwidth, large number of channels, and high resolution, and is compact, lightweight, and energy efficient. These factors become very important as heterodyne receivers are developed for ever higher frequencies.

The thrust of receiver development is towards high frequency heterodyne systems, particularly in the millimeter, submillimeter, far infrared, and 10 μm spectral ranges. The motivation for this development comes from the need to determine the chemical composition of the interstellar medium and to achieve a thorough understanding of the excitation of interstellar molecules. The most important constituent of the interstellar medium after molecular hydrogen is carbon monoxide (CO) which has its rotational transitions at 115.3, 230.5, 345.8, and 691 GHz, etc., while hydroxyl (OH), an equally important constituent, has its ground-state rotational transition at 2508 GHz. In general, the lighter the molecule, the higher its ground-state transitions which fall often above currently accessible receiver ranges.

The LO, LO coupler, and mixer present difficult challenges as the range of the receiver is extended towards higher frequencies. Klystrons which are already expensive and short-lived in the 3 millimeter range become impractical at shorter wavelengths forcing development of frequency doubling or tripling mixers utilizing cheaper and more robust lower frequency LOs. In the 10 μm band the CO_2 gas laser-heterodyne system, with its various isotopic mixtures, can cover only about 10% of the spectral band due to the limited tunability of the gas laser transition and limited bandwidth response of mixers. Work is presently underway in this laboratory to construct a diode heterodyne system which will have a 2 μm continuously tunable range. RF tuned cavity LO couplers which work well in the 3 millimeter wavelength become highly lossy in the millimeter range forcing development of quasi-optical techniques to combine signal and LO efficiently to the mixer. Mixer development is also continuing to extend usable range of mixers into the far infrared and 10 μm bands. Recently we have used a 25 watt CO_2 10 μm laser to pump a formic acid 430 μm laser in an experiment on the 3-meter IRTF in Mauna Kea to detect and map the 691 GHz, $J = 6 \rightarrow 5$ transition of CO in Orion and other galactic sources.

Common to all heterodyne systems is an RF spectrometer (back end) which recovers the power spectrum of the signal from the IF. Traditionally the spectrometer is a bank of discrete adjacently tuned RF filters or a digital auto-correlator. Most radio telescopes have 256 or 512 channels at 1 MHz or 0.25 MHz resolution. Heterodyne receivers in the millimeter, submillimeter, or far infrared will be operated from remote mountain, airborne, balloon-borne, or spaceborne platforms to avoid the severe atmospheric attenuation in the 20 μm to 1 mm range. This will place severe demands of size, weight, and energy usage on the back end which RF filters may not be able to satisfy.

Recent developments in acousto-optic techniques and in photodetector arrays have made feasible a new type of RF spectrometer offering the advantages of wide bandwidth, high resolution, large number of channels in compact, lightweight, energy efficient, and relatively low cost systems. Such a system employs an acousto-optic diffraction cell which serves the key role of converting RF signals to ultrasonic traveling waves modulating the optical index of the cell. The cell is illuminated across its aperture by a monochromatic laser beam. A fraction of the light is diffracted by the acoustic waves; the angle of diffraction is determined by the frequency while the intensity of the diffracted light is proportional to the power of the input RF signal. (The major portion of the laser beam at zero order passes through the cell undeflected.) A focusing lens follows the cell and essentially performs a Fourier transform of the RF signal into a far-field intensity pattern. The output intensity distribution is typically received by a linear array of photodetectors whose output is the RF power spectrum we seek. The advantage of an AOS is due to the simplicity arising from the small number of components needed to build up the system.

Currently there is intense commercial interest in applying acousto-optic techniques to electronic warfare (EW), electronic counter measures (ECM) and electronic support systems (ESM) which is spurring rapid technical advancements in the field. The Air Force and Navy, in a joint effort, are funding research at GTE-Sylvania, ATI-Itek, ESL, Teledyne-MEC, and Rockwell International in bulk Bragg cell AOS components. One benefit from this research is the commercial availability of 1 GHz bandwidth, 1 MHz resolution Bragg cells. The Air Force and the Navy have also funded Hughes Research and Westinghouse in an effort to further miniaturize an acousto-optic system by integrating all its components on a single substrate. These recent developments have made acousto-optic techniques increasingly attractive for use in astronomical applications.

We are currently both experimenting with a bulk and awaiting delivery of an integrated AOS device for astronomical research here at Goddard. The GSFC prototype AOS uses a discrete bulk acoustic wave Itek/Applied Technology Bragg cell with 300 MHz bandwidth (specified) 0.67 MHz resolution (500 resolution elements), 5 mW Spectraphysics Model 120 helium-neon laser with aluminum optical bench components assembled from Klinger Scientific. The length of the stainless-steel rods is 1 meter, and although the optical layout is fairly compact, no attempt was made to optimize the unit for size. The laser is located on the bottom of a two tier construction, with the beam directed to the upper level by guiding mirrors. A beam expander (16X) is placed in front of the Itek Bragg cell with the cell mounted on X-Z position translators. An aberration-minimized biconvex lens 80 mm in diameter with a 47 cm focal length follows the Bragg cell. The diffracted light is then guided downward by a mirror flat, and the output light is detected by a 1024 element CCPD Reticon array mounted on precision rotation and translation stages. The interlocking construction of the Klinger assembly gives the system very good rigidity and once the optical path has been aligned, retains its alignment even after movement of the assembly as a unit. Although no attempt has been made to temperature control the assembly the AOS appears to drift less than 1 channel over a 24-hour period.

The Itek/Applied Technology Bragg cell is made from a 1 cm optical aperture LiNbO$_3$ crystal with a specified 300 MHz bandwidth centered at 450 MHz. The interaction time is 1.5 ms with a time-bandwidth factor of 450. The diffraction efficiency of the cell is reported to be 7%/watt of RF power. Preliminary results of the GSFC AOS are given elsewhere[1].

An exciting prospect is the use of integrated optics in further miniaturizing an acousto-optic spectrometer. Recently, Westinghouse, under NRL contract and GSFC participation, has constructed a 400 MHz bandwidth, 140 element integrated AOS[2]. This entire RF spectrometer has dimensions of 7X3 cm. The spectrum analyzer, shown schematically Fig. 1, consists of a laser, a diffraction-limited geodesic collimating lens, a surface acoustic wave (SAW) transducer array, a second diffraction-limited geodesic lens that is utilized as a transform lens, and a 140 element photodiode array. A Te polarized 100-mW He-Ne laser operating at 0.6328 mm was end-fire coupled into the spectrum analyzer for preliminary testing. Efforts are now underway by Westinghouse to butt couple a diode laser source to the substrate.

The spectrum analyzer was fabricated on X-cut LiNbO$_3$ with the c axis parallel to the acoustic propagation. The optical waveguide was formed by diffusing 180 A of titanium into LiNbO$_3$ to obtain a tightly confined optical beam. The photosensor array was butt coupled at a 45° angle to the waveguide edge of the LiNbO$_3$ substrate in an optically polished Cer-Vit mounting block. This angular mounting minimizes spurious signals due to reflected light from the surface of the photodiode array and waveguide substrate.

The laser beam is coupled into the waveguide and collimated by the first geodesic lens. The SAW spatially modulates the guided optical beam. The deflected beam is focused onto the output edge of the optical waveguide. The butt-coupled photodiode array then detects the deflected beam. The deflection angle is proportional to the frequency of the RF signal. Therefore, the frequency of an incoming signal can be determined from the position of the focused beam on the detector array.

The geodesic lenses are circularly symmetric aspherical depressions in the surface of the waveguide and are fabricated by single-point diamond turning. The advantages of using single-point diamond turning for fabricating geodesic lenses are numerous. First, the lens can be precisely located on the substrate, which allows the focal plane of the lens to be placed at a predetermined position. Second, the lens profile can be cut to tolerances of better than 0.5 mm, which gives diffraction-limited performance and predetermined focal lengths. Finally, surfaces of the lenses can be cut very smoothly, thus eliminating the need for much post-machining surface polishing. We have measured the performance of diamond machined lenses and have found them to be diffraction-limited with an insertion loss of 2 dB, and to have a focal plane which is within 12 mm of the previously polished waveguide edge.

The detector array is a self-scanned photodiode array consisting of 150 photodiode pixels with a 12-mm pitch. The maximum shift register clockrate is 5 MHz, resulting in an access time of 2 msec. The maximum measured output voltage was 1.0 volt while the minimum measured output voltage was 90 mvolt, which yields a dynamic range of 40 dB. The next nearest neighbor crosstalk was down by 15 dB. A single zeroth-order beam sensing photodiode is included on the 6.4-mm square chip. The array is divided into 7 pixel groups of 20 pixels each. Each group is addressed in parallel by a 10-stage dynamic PMOS shift register. Each group has two electrometer output circuit channels; one handles odd pixels, while the other handles the even pixels for each clock cycle of the PMOS shift register. This scheme provides fourteen parallel samples every clock cycle.

The surface acoustic wave transducers are a two-element tilted array designed to have a combined 400-MHz bandwidth centered at 600 MHz. The operation of the spectrum analyzer and the frequency range over which the transducers operate are shown in Fig. 2. A separate measurement of the transducer performance provided 3-dB bandwidth of 400 MHz and a 5% deflection efficiency with an RF power of 60 mW.

In summary, this laboratory is in the process of constructing and evaluating bulk AOS devices with 300 MHz and 1 GHz bandwidths for use in the back end of high frequency heterodyne receivers for use in astronomical research. In addition, we are in the process of obtaining and will evaluate the feasibility of using integrated AOS devices as potential back ends for future balloon-borne or space experiments using heterodyne receivers.

References

1. SPIE Proceedings 1980 International Optical Computing Conference, Volume 231, p. 30, Chin, Buhl and Florez.

2. Applied Optics, Volume 19, No. 18, September 15, 1980, Mergerian, Malarkey, Pautienus, Bradley, Marx, Hutchenson, and Kellner.

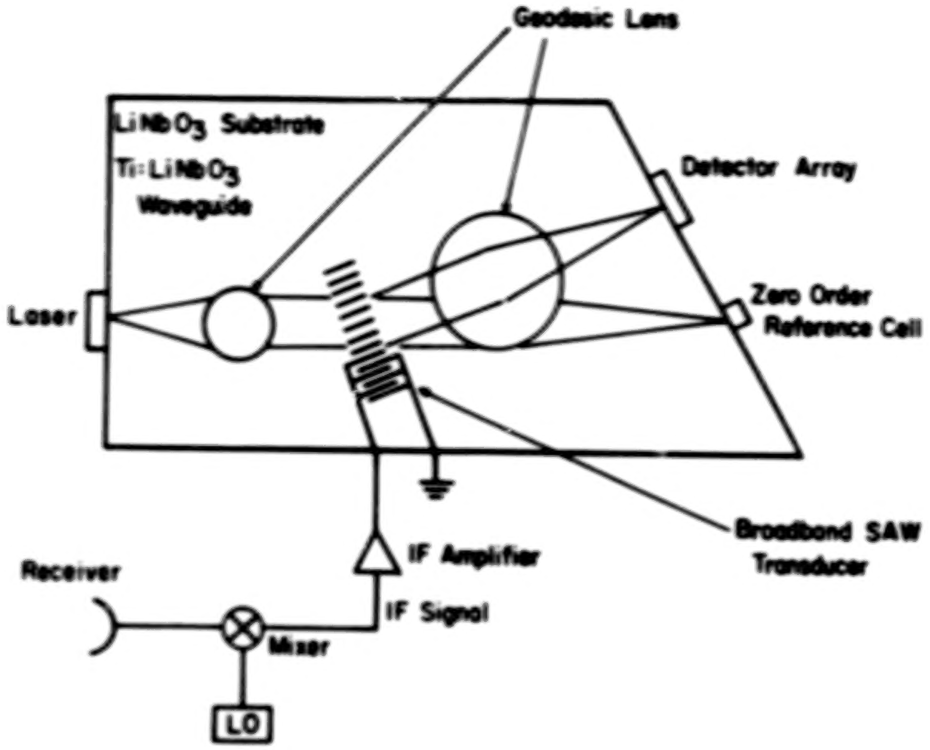

Geodesic Lens
LiNbO$_3$ Substrate
Ti:LiNbO$_3$ Waveguide
Detector Array
Laser
Zero Order Reference Cell
Broadband SAW Transducer
IF Amplifier
Receiver
IF Signal
Mixer
LO

Figure 1.- Schematic of an integrated acousto-optics spectrometer.

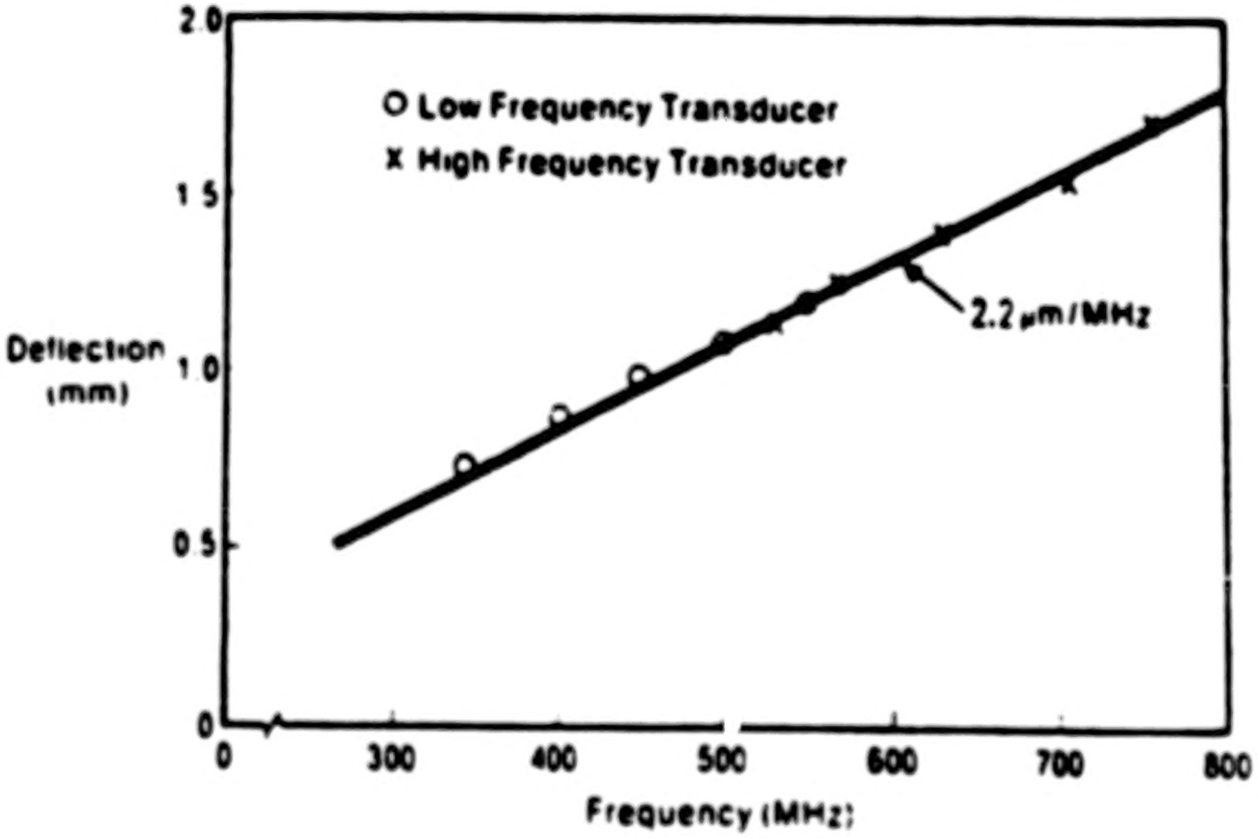


Figure 2.- Deflection of optical guided wave as a function of surface acoustic wave frequency.

ENERGY NORMALIZATION OF TV VIEWED OPTICAL CORRELATION
(Automated Correlation Plane Analyzer for an Optical Processor)

Alex Grumet
Grumman Aerospace Corporation
Bethpage, New York

SUMMARY

An automatic digital correlation plane processor is described that permits rapid location and identification of up to 99 different normalized autocorrelations. Up to 99 unnormalized correlations are acquired in a single TV frame of 1/30 sec. Analog preprocessing circuits permit digital conversion and RAM storage of those video signals with the correct amplitude, pulse width, rising slope, and falling slope. TV synchronized addressing of 3 RAMs permits on-line storage of: 1) the maximum unnormalized amplitude, 2) the image x location, and 3) the image y location of the output of each of up to 99 matched filters. A fourth RAM stores all normalized correlations.

Only the maximum correlation for each of 99 matched filters is determined in one TV frame of 1/30 sec. Successive TV frames look for the next largest correlation output for each optical matched filter.

The silhouette of each image stored in the multiple optical matched filter is also stored in a pair of ROMs that represent the left and right side of the silhouette. Input image energy is collected through an optical high-pass filter of the same bandwidth as the corresponding optical matched filter by a gated, wideband video integrator. The image energy signal is routed to a ROM reciprocal look-up table. Normalization (division by image energy) is accomplished in 1 μsec by a dedicated digital multiplier. A partial out-of-field signal is available to abort normalization and to indicate the approach of loss of track.

INTRODUCTION

The optical matched filter (MF) (reference 1) has not found wide application because acceptable false alarm rates have not been realized by observation of the amplitude of the correlation peaks. When the illumination level associated with an unmatched input image greatly exceeds that of the matched image of comparable size, the MF generally yields a false alarm if MF output correlation peak is the only measure of acceptance. However, if we could divide the above autocorrelation and cross-correlation output peaks each by their corresponding input image energies, the resulting normalized correlations would then be independent of input image energy, and for this situation, the normalized autocorrelation peak would be larger.

The normalization approach described below does normalize the autocorrelation peak but yields an approximate normalization of the cross correlation peak. We, therefore, expect the normalized autocorrelation not to always be larger, but to always be unity, whereas the approximate normalized cross correlation will rarely be unity. The deviation from unity will be used as a measure of the uncertainty of a correct image match. To assist in reducing false alarms, other discriminants such as unnormalized correlation pulse width, rising and falling slopes, as well as unnormalized pulse amplitude are an integral part of the system.

PHENOMENOLOGIC MATCHED FILTER REVIEW

Because we (1) exploit the automatic tracking feature of the MF, (2) use multiple MFs, and (3) employ band-limited MF's to normalize the correlation outputs, a brief description of the MF is in order.

Figure 1 is one of the many possible MF configurations. The back focal plane of the transform lens contains the on-axis, two-dimensional Fourier transform $G(f_x, f_y)$ of $g(x,y)$, the amplitude of the matched input image. When we signify the input image we refer to the image of the object of interest in the input scene. During fabrication it will be the only image in the scene. The spectrum $G(f_x, f_y)$ is complex and two-dimensional and its phase can be recorded on high resolution spectroscopic film by use of a second phase-locked collimated laser beam at angle θ to the signal beam. The resulting stationary interference pattern shown magnified in figure 1 contains a grating, or "subcarrier," whose spacing depends on the angle between the signal and reference beams. Superimposed on the subcarrier grating are the very slight perturbations of the phase contributions of the Fourier transform of the input image. On playback with the reference beam removed and with the exposed and developed MF inserted in the identical fabrication position, the conjugate first order of the MF subcarrier is used as shown in figure 1. In figure 2 we fold the arrangement of figure 1 in line to demonstrate several points. First, for the matched input image the light exiting the MF is $G(f_x, f_y)\, G^*(f_x, f_y)$, where $G(f_x, f_y)$ is the amplitude of the input matched spectrum and $G^*(f_x, f_y)$ is the amplitude transmittance of the matched filter. Therefore, all rays exiting the MF, regardless of their input direction (phase), are parallel. The consequence of this fact is that they focus to a point at the correlation output plane. The focused point in the correlation plane is the summation of all spatial frequencies and, therefore, all the input image energy less system transmission losses. This is not true for an unmatched input image and the cross-correlation peak will be reduced as well as broadened. For a band-limited MF with the low spatial frequencies shown in the magnified MF of figure 1 saturated (overexposed), the autocorrelation function becomes a sharp point, as shown in figure 3(b) and compared to that of figure 3(a) for the full spectrum MF. In practice, the photographic MF plate has a dynamic range of the order of 1000:1 and the energy spectrum of most objects has a dynamic range of the order of 10^6:1, so that the full spectrum autocorrelation of figure 3(a) is not realizable. Second, several MFs can be processed simultaneously through the use of a multiple holographic lens array that is equivalent to several overlapping lenses as shown in figure 2(b). Third, figure 2(a) indicates a transverse displacement of the input image, and, for $F_o = F_1$, the correlation will experience the same displacement with an inversion in the x and y directions. Therefore, input image position in the input plane can be determined by observing the position of the correlation output. By making $F_1 < F_o$, the mapping of the input image plane into the output correlation plane can be reduced to fit the output of several filters into the TV camera tube viewing aperture.

NORMALIZATION APPROACH

Parseval's theorem equates image integrated energy to its spectral

$$\int_y \int_x g^2(x,y)\,dx\,dy = \int_{F_y} \int_{F_x} |G(f_x,f_y)|^2 df_x df_y$$

integrated energy. The left side of the equation, image energy, is measured from the input image and provides the denominator for the normalized correlation. The right side, spectral energy, is measured by the autocorrelation peak to provide the numerator. Provision is incorporated to collect image energy in the same restricted band of the MF. Image energy band limiting is accomplished by selecting one of the 100 holographic lenses for imaging. Instead of a MF at this beam we provide a high-pass or bandpass imaging filter of the same bandwidth as the MF, in the MF plane. The reference beam, on fabrication, is used to direct the band limited image on playback to an imaging lens to provide a different image plane scaling and a different TV pickup other than that used for the correlation plane. In a feasibility experiment now being assembled, a 2 x 2 holo-lens will provide three MF and one imaging lens that will share the same inverse transform lens for MF, as well as imaging, and will share the same TV camera for correlation and imaging.

To collect the image energy, the TV video is integrated only during the time the desired image is scanned. Detectors are square-law devices, and the TV video voltage is proportional to energy. An earlier version of this concept restricted video integration to a rectangle that was centered on the desired image. The x-y RAMs provide the integrating window position over the image. The particular MF, as determined by the RAM address, determines the size of the rectangle. In this manner the matched image will always fall within the rectangle window even though an unmatched image may fall partially outside. However, since the input image and its background are a high frequency version of the actual input scene, we can expect a non-negligible contribution by the background around the matched image within the rectangle. To avoid this situation, and to guarantee that the normalization of the autocorrelation will be accurate and free of background, the shape of the rectangle was altered to that of the silhouette of the matched image.

Two ROMs are used to store the silhouettes of all matched images. We will refer to one as the START ROM and the other as the STOP ROM. The silhouette in each interlaced field of a TV frame is located at contiguous addresses in the START and STOP ROMs. For silhouettes where a TV scan line leaves and reenters, an additional pair of ROMs will be required to avoid integrating background energy. A TV frame of 1/30 sec is required to collect the image energy of each selected unnormalized correlation. Preprocessing circuitry limits the number of unnormalized correlations to be normalized.

APPLICATION

The need for normalized correlations can be appreciated if we have to select or detect a possible autocorrelation out of thousands of correlations in a few seconds. This was the motivation for the development of the system to be described.

The OMFIC (Optical Matched Filter Image Correlator) is described in reference 2 and is presented in schematic form in figure 4. The system is capable of processing an input image through 100 optical matched filters simultaneously. Therefore, an image of an input scene containing 20 objects will yield 2,000 correlations in 2 or 3 ns at the OMFIC output plane. The output correlation plane is scanned by a conventional TV camera and all 2,000 correlations can be viewed on a TV monitor.

The problem addressed was to select any autocorrelations out of the array of correlations, and identify the image (associated matched filter) and the image x-y coordinates in a few seconds. Correlation normalization was one of the discriminators used to reduce false alarms. If the matched images were trucks parked side by side, the resulting autocorrelations could be a few microseconds apart, and this ruled out the use of microprocessors to process the TV video. TTL with 20 ns switching time are adequate and, therefore, the bulk of the system described below consists of integrated TTL circuits. A μP is used to provide slow information in a few 100 μs time frame and to supervise overall operations, conduct searches, and provide displays and printout messages.

SYSTEM DESCRIPTION

Figure 5 is an overall system block diagram. The TV video is preprocessed so that only video pulses of the correct amplitude, of the correct pulse width, and with the correct rising and falling slopes will be accepted to be converted to 8-bit digital data in a 600 ns A/D converter. Figure 6 shows the preprocessing gate timing. The unnormalized correlations are stored in a 256 x 8 bit RAM. There are three additional RAMs, all 256 x 8 bits, and two of these are synchronously addressed with the TV raster sweep, along with the unnormalized correlation RAM as described below. Two of the RAMs store image x-y coordinates, and the fourth RAM stores the normalized correlations.

Two timing counters accurately address the RAMs and locate the position of all correlations to the required precision on the TV raster. The horizontal counter consists of a digital phase-locked-loop (PLL) and provides the four least significant bits (LSB) for the RAM 8-bit address bus. The vertical counter is essentially a TV line counter providing the four most significant bits (MSB) for the RAM address. The H and V counts can be hardware or software altered to agree with the number and configuration of the matched filter array. A test pattern can be displayed on a TV monitor for system setup and assistance in matched filter fabrication output addressing. Such a pattern is shown in figure 7 for a 10 x 10 matched filter array. Each of the 100 boxes in the checkerboard test pattern represents a RAM address. The bright spot in the upper left-hand corner of each box (more visible in figure 3(b)) is the RAM erasure pulse of 100 ns duration of every other TV field. Since the RAMs are synchronously addressed with the TV raster, the data at any one address can be displayed on the TV monitor. The MSB of the data bus was connected to the TV monitor video input to demonstrate that the correlation peak was stored at the correct address. To ensure that only the peak of the video pulse is stored in each address, an 8-bit digital comparator tests the incoming 8-bit video peak with the RAM stored value and will dump the stored value if it is smaller. In this manner, only one correlation peak is stored in each address. If several correlation pulses are present for a particular MF (in the same box of the checkerboard), the stored value will be inhibited by the μP on the next TV frame after normalization. The next largest correlation peak of the particular filter will then be stored. After several TV frames all correlation peaks associated with each matched flter will be processed. The 8-bit comparator output pulse triggers a write pulse for the unnormalized and the x and y RAMs. Therefore, the maximum unnormalized correlation peak of each filter is stored in a RAM address assigned to that filter and the x-y coordinates associated with the correlation are stored at the same address in their respective x and y RAMs. At the end of each TV frame the RAMs are erased and the process is repeated.

During every other TV vertical blanking interval (once per TV frame) of approximately 1.7 ms all three RAMs (unnormalized correlation, x and y coordinates), together with the normalized correlation RAM, are scanned very rapidly by switching in a high-speed addressing counter. At each address the contents of the unnormalized and the normormalized RAMs are interrogated for any content. If no normalized number is stored where an unnormalized number has been found, then the normalization process will be initiated during the next TV frame, and the first step in the normalization process is the storage of 1) the RAM address, 2) x count, 3) y count, and 4) the unnormalized correlation.

In figure 8, we define the image height and location, referenced to its location, as stored in the x and y RAMs. The left shaded side of the image, designated SR, is stored in the START ROM at contiguous addresses for successive TV lines in a field, and the interlaced field is stored at another group of contiguous addresses. The same image is stored at the same addresses in the STOP ROM for the right edge of the image, designated SP. When the horizont[illegible] and, therefore, the x counter of the timing chain is equal to the value st[illegible] START ROM, the associated comparator output will switch on a high-speed, w[illegible] integrator, triggered by SR, and, in a similar manner, the STOP ROM will [illegible] integrator off, triggered by SP. The height, TL + BL, of the image trigger[illegible]ght FF to gate in only the correct SR and SP ROM addresses. To correctly address the START and STOP ROMs for: a) the correct image, b) correct field, and c) correct positon in x and y requires offsets on all position data. Some of these are slow (TV field intervals) and are provided by the microprocessor. The 8-bit counts that are required in less than a TV line interval of 63.5 μs are hard wired.

The RAM address identifies the MF and, hence, the image. The location of this image, in the START and STOP ROM, must be accessed as the TV sweeps through the image. The value $Y_o - T_L$ of figure 8 defines the image top line, and the μP provides the offset for the corresponding ROM location of this image. In this manner, the vertical timing chain line-counter is offset to correctly address the ROMs as the image is scanned. At the beginning of each TV scan line, the ROM stored value is subtracted from the X_o stored value to provide a count for the leading edge of the image (X_o - SR). An 8-bit comparison of this value with the horizontal PLL timing chain identifies the left edge of the image to start the video integrator. In a similar manner (X_o + SP) will stop the integrator for that particular TV line. After a complete TV frame, the integrator output is converted to 8 bits and a reciprocal ROM provides the denominator to a fast (1 μs) 8 bit by 8 bit multiplier with the unnormalized correlation. Storage of this value in the normal correlation RAM completes the normalization cycle. During the TV frame that normalization is underway, the X-Y RAMs, as well as the unnormalized RAM, continue to be updated as before. The μP reads the normalized RAM and checks for deviation from unity. An appropriate video overlay circle or square can encircle the correlation peak, and the blinking rate of the overlay can indicate deviation from unity.

The block diagram of figure 9 indicates how the operations described above are realized in a straightforward fashion. To determine if the image is partially cut off by the raster boarder, we evaluate four quantities. If $(Y_o - T_L) < 0$, then the image is partially off the top of the screen. Similarly, if $(Y_o + B_L) > 244/K$ (where K is the number of rows of MF, and, therefore, correlation address rows), the image is partially off the bottom of the screen. If (X_o - SR) < 0, then the image is partially off the left side of the raster, and if (X_o + SP) > 100 (for a 100 count raster), then the image is partially off the right side of the screen. In all four of these conditions, the normalization will be aborted and an appropriate message will be indicated by the μP. This same information could be used for an automatic tracker associated with the OMFIC that is about to lose its target and the direction to move to maintain the track is indicated.

NORMALIZATION OF CROSS CORRELATIONS

We can apply Parseval's theorem to the cross correlation situation

$$\int_y \int_x g_2(x,y) g_1(x,y) dx dy = \int_{F_y} \int_{F_x} G_2(f_x,f_y) G_1^*(f_x,f_y) df_x df_y$$

where $g_2(x,y)$ is an unmatched input image amplitude and $g_1(x,y)$ is the matched image. However, we do not have a single peak in the correlation plane that represents the integrated cross product because G_2 and G_1 do not have conjugate spectra, and the spectra do not have the same spatial frequencies. A cross correlation peak need not be an even function (as it is for an autocorrelation peak), and its location can be located almost anywhere within the input image if we map it back to the input plane. The ROM stored silhouette location is positioned by the x-y RAM stored correlation peak as indicated in figure 8. There are three different situations that we can encounter: 1) the extent of g_2 > the extent of g_1, 2) the extent of g_2 < the extent of g_1, and 3) g_1 and g_2 do not completely overlap. In case (1) where the silhouette g_1 falls entirely within g_2, the input image, we do collect less than the full input image energy, and the denomination of the normalization is smaller than for a perfect g_2 silhouette. It is, therefore, possible to have a normalized cross correlation greater than unity. In case (2) we can have a smaller than unity normalization; however, we do collect background, and this can alter the normalization. The same is true for case (3).

Therefore, although we can make a definitive statement that the normalized autocorrelation is unity, we can expect the normalized cross correlation to vary considerably with some slight possibility of a value of unity.

No effort was made to exploit the even symmetry of the autocorrelation peak in the preprocessor. If required, this extra feature can be implemented in a straighforward manner.

CONCLUSIONS

The automatic correlation plane processor described here can rapidly acquire, identify, and locate the autocorrelation outputs of a bank of multiple optical matched filters. The ROM stored digital silhouette of each image associated with each matched filter allows TV video to be used to collect image energy to provide accurate normalization of autocorrelations. The resulting normalized autocorrelations are, therefore, independent of the illumination of the matched input. Deviation from unity of a normalized correlation can be used as a confidence measure of correct image identification.

REFERENCES

1. Vander Lugt, A. B.: Signal Detection by Complex Spatial Filtering, IEEE Transactions on Information Theory, April 1964, pp 139-145. A more complete treatment is given in AD 411473, a University of Michigan Report 4594-22-T, July 1963, issued by the Radar Laboratory of the Institute of Science and Technology.

2. Grumet, A. and Reich, A.: Optical Matched Filtering for Missile Guidance, Final Report Contract N00019-73-C-0303, NASC, April 1974.

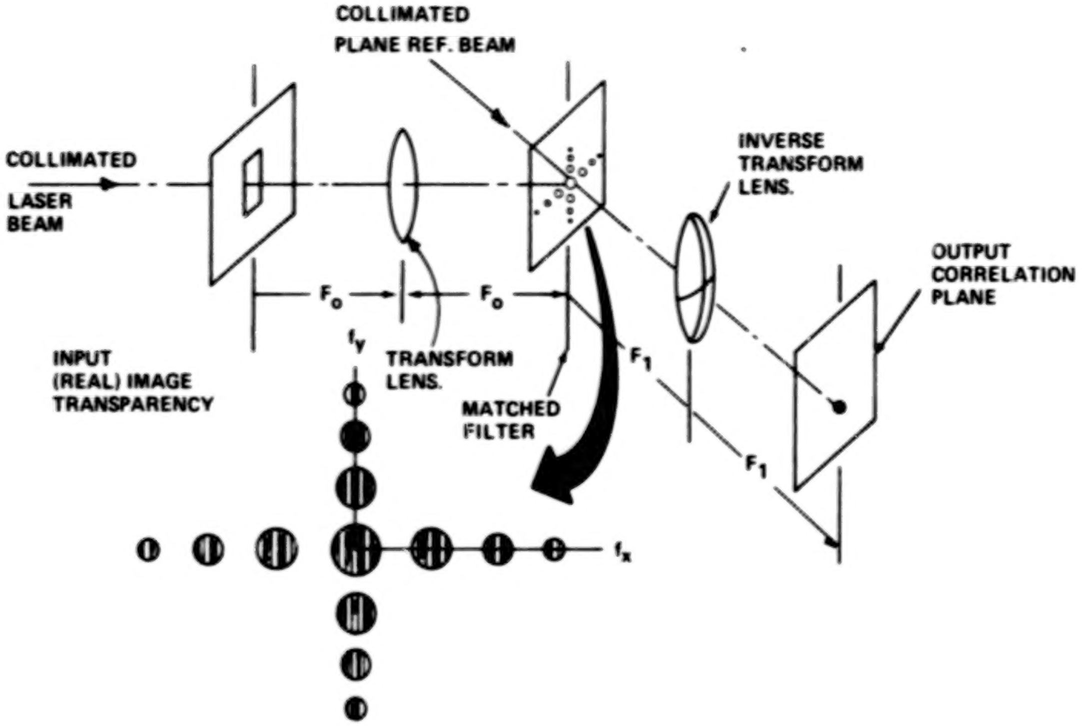

COLLIMATED
PLANE REF. BEAM
COLLIMATED
LASER
BEAM
INVERSE
TRANSFORM
LENS.
OUTPUT
CORRELATION
PLANE
F₀
F₀
f_y
F₁
F₁
INPUT
(REAL) IMAGE
TRANSPARENCY
TRANSFORM
LENS.
MATCHED
FILTER
f_x

Figure 1.- Matched filter configuration.

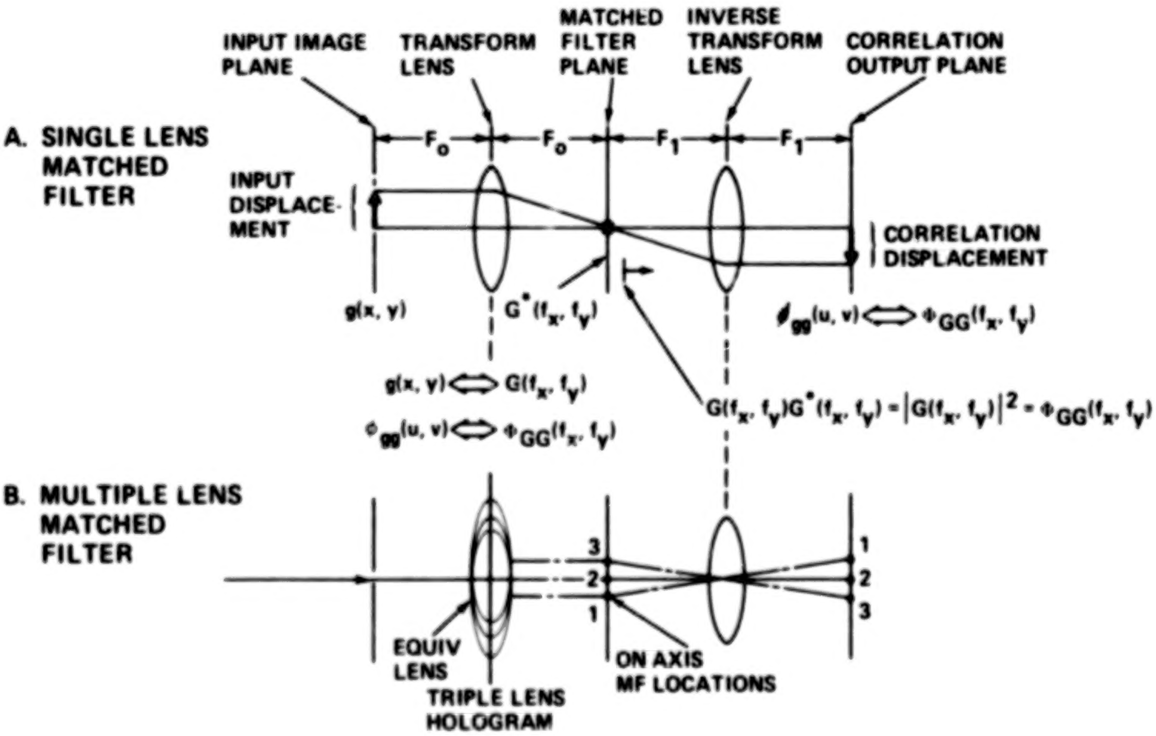

INPUT IMAGE
PLANE
TRANSFORM
LENS
MATCHED
FILTER
PLANE
INVERSE
TRANSFORM
LENS
CORRELATION
OUTPUT PLANE
A. SINGLE LENS
MATCHED
FILTER
F₀
F₀
F₁
F₁
INPUT
DISPLACE-
MENT
CORRELATION
DISPLACEMENT
g(x, y)
G*(f_x, f_y)
φ_gg(u, v) ⟺ Φ_GG(f_x, f_y)
g(x, y) ⟺ G(f_x, f_y)
φ_gg(u, v) ⟺ Φ_GG(f_x, f_y)
G(f_x, f_y)G*(f_x, f_y) = |G(f_x, f_y)|² = Φ_GG(f_x, f_y)
B. MULTIPLE LENS
MATCHED
FILTER
3
2
1
1
2
3
EQUIV
LENS
TRIPLE LENS
HOLOGRAM
ON AXIS
MF LOCATIONS

Figure 2.- Some aspects of matched filters.

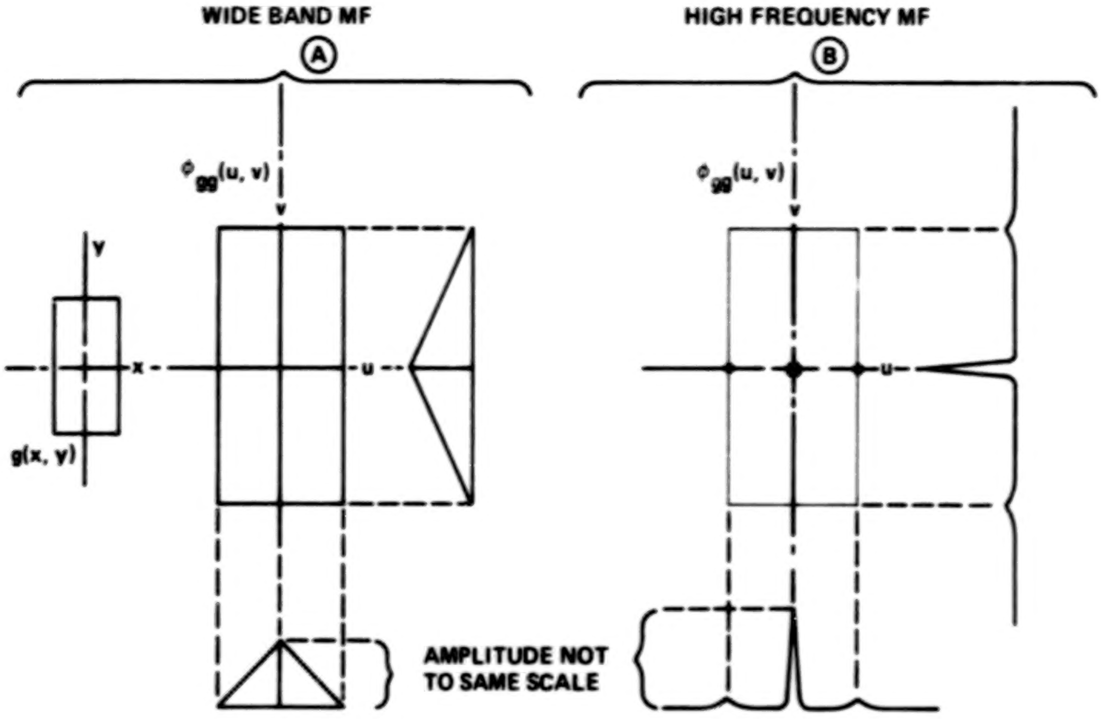

WIDE BAND MF
Ⓐ
HIGH FREQUENCY MF
Ⓑ
$\phi_{gg}(u, v)$
$\phi_{gg}(u, v)$
y
x
u
u
g(x, y)
AMPLITUDE NOT
TO SAME SCALE

Figure 3.- Wide band and high frequency autocorrelation functions.

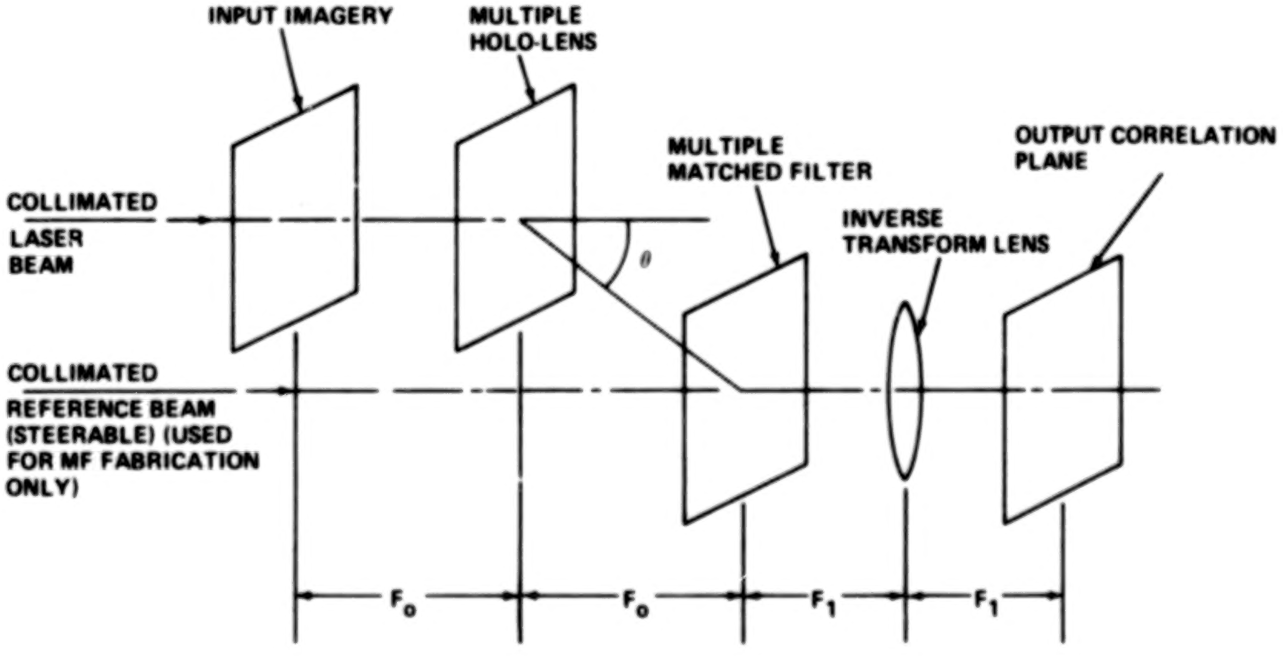

INPUT IMAGERY
MULTIPLE
HOLO-LENS
MULTIPLE
MATCHED FILTER
OUTPUT CORRELATION
PLANE
INVERSE
TRANSFORM LENS
COLLIMATED
LASER
BEAM
θ
COLLIMATED
REFERENCE BEAM
(STEERABLE) (USED
FOR MF FABRICATION
ONLY)
F_o
F_o
F_1
F_1

Figure 4.- Optical matched filter image correlator (OMFIC).

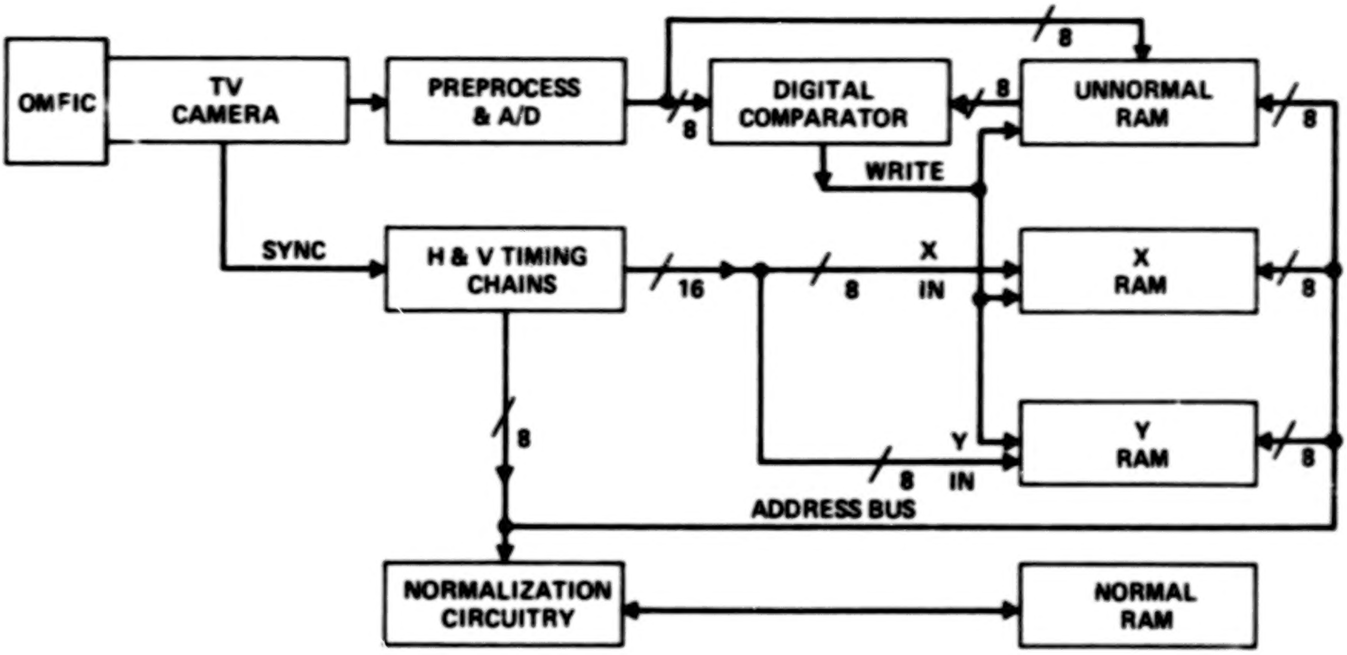

OMFIC
TV CAMERA
PREPROCESS & A/D
DIGITAL COMPARATOR
UNNORMAL RAM
8
8
8
8
WRITE
SYNC
H & V TIMING CHAINS
16
8
X IN
X RAM
8
Y RAM
8
Y IN
8
ADDRESS BUS
NORMALIZATION CIRCUITRY
NORMAL RAM

Figure 5.- Overall system block diagram.

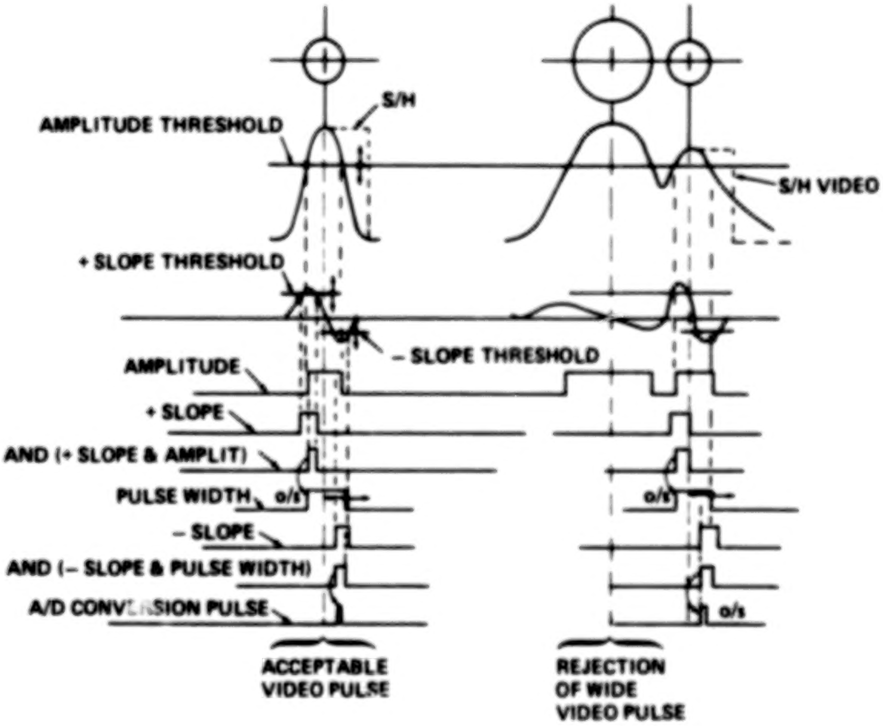
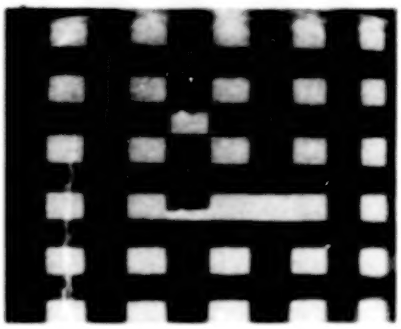
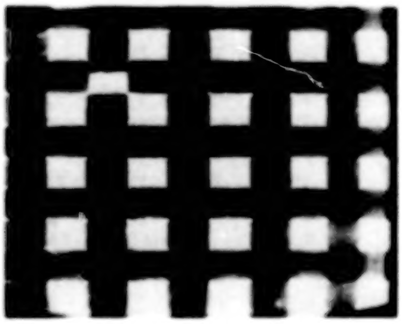

S/H
AMPLITUDE THRESHOLD
S/H VIDEO
+ SLOPE THRESHOLD
– SLOPE THRESHOLD
AMPLITUDE
+ SLOPE
AND (+ SLOPE & AMPLIT)
PULSE WIDTH
o/s
– SLOPE
AND (– SLOPE & PULSE WIDTH)
A/D CONVERSION PULSE
o/s
ACCEPTABLE VIDEO PULSE
REJECTION OF WIDE VIDEO PULSE

Figure 6.- Video preprocessing waveshapes (condition for A/D conversion).

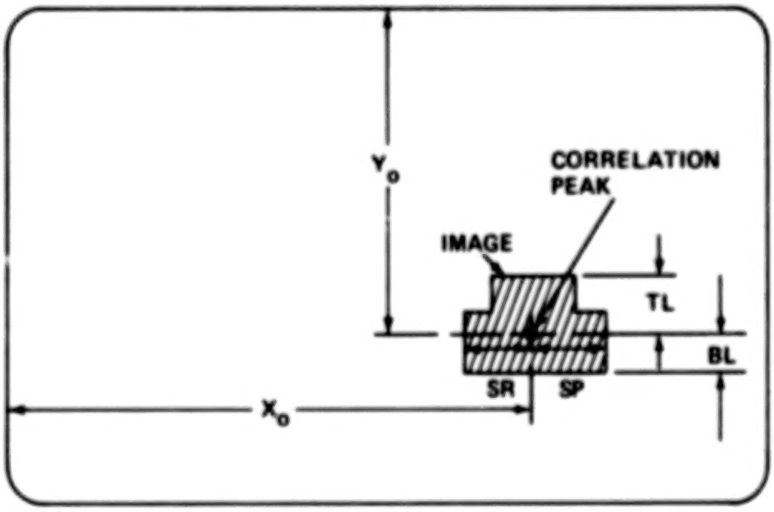


(a) THREE CORRELATIONS PRESENT



(b) ONE CORRELATION PRESENT

Figure 7.- Ten by ten test pattern.


Y_o
CORRELATION PEAK
IMAGE
TL
BL
SR
SP
X_o

X_o = x COORDINATE OF THE CORRELATION STORED IN X–RAM
Y_o = y COORDINATE OF THE CORRELATION STORED IN Y–RAM
TL = TOP LINE OF IMAGE
BL = BOTTOM LINE OF IMAGE
SR = 8-BIT NUMBER STORED IN START ROM
SP = 8-BIT NUMBER STORED IN STOP ROM

Figure 8.- Definition of terms.

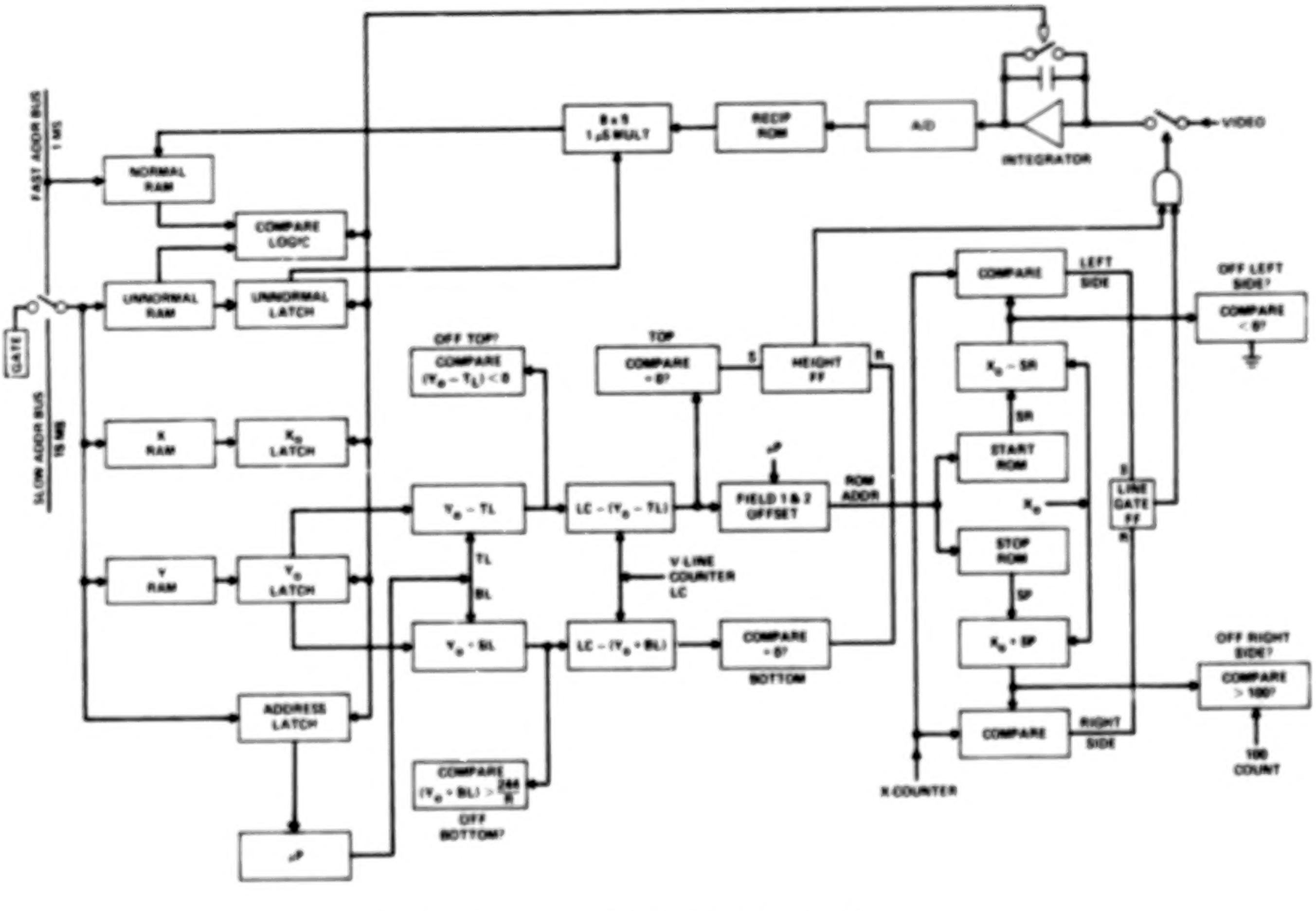

FAST ADDR BUS
1 MS
SLOW ADDR BUS
15 MS
GATE
NORMAL RAM
COMPARE LOGIC
UNNORMAL RAM
UNNORMAL LATCH
X RAM
X_o LATCH
Y RAM
Y_o LATCH
ADDRESS LATCH
µP
8 x 8 1 µS MULT
RECIP ROM
A/D
INTEGRATOR
VIDEO
OFF TOP?
COMPARE $(Y_o - T_L) < 0$
TOP
COMPARE = 0?
S
HEIGHT FF
R
Y_o – TL
TL
BL
Y_o + BL
LC – (Y_o – TL)
LC – (Y_o + BL)
V LINE COUNTER LC
µP
FIELD 1 & 2 OFFSET
ROM ADDR
COMPARE = 0?
BOTTOM
COMPARE $(Y_o + BL) > \frac{244}{R}$
OFF BOTTOM?
COMPARE
LEFT SIDE
OFF LEFT SIDE?
COMPARE < 0?
X_o – SR
SR
START ROM
X_o
STOP ROM
SP
X_o + SP
S
LINE GATE FF
R
OFF RIGHT SIDE?
COMPARE > 100?
100 COUNT
COMPARE
RIGHT SIDE
X-COUNTER

Figure 9.- Normalization block diagram.

ITERATIVE OPTICAL VECTOR-MATRIX PROCESSORS

(SURVEY OF SELECTED ACHIEVABLE OPERATIONS)

David Casasent and Charles Neuman
Carnegie-Mellon University
Department of Electrical Engineering
Pittsburgh, Pennsylvania 15213

ABSTRACT

An iterative optical vector-matrix multiplier with a microprocessor-controlled feedback loop is capable of performing a wealth of diverse operations. In this paper, we survey and describe many of these operations to demonstrate the versatility and flexibility of this class of optical processor and its use in diverse applications.

1. INTRODUCTION

The optical vector-matrix multiplier [1] is a general purpose optical processor. The addition of a microprocessor-controlled feedback loop results in an even more general purpose and far more powerful optical processor [2-6]. In this paper, we survey and describe a selected set of the operations achievable on such a processor.

In Section 2, a general description of the system is advanced. This is followed in Section 3 by a description of how bipolar and complex-valued data are handled on the system and how the convergence of the iterative algorithm is insured. We then address in Section 4 its use in the solution of linear difference and differential equations and linear algebraic equations. In Section 5, we consider application of this processor for the solution of the least-squares problem. In Section 6, we address its use for deconvolution and for the computation of the eigenvalues and eigenvectors of a matrix. In Section 7, our attention turns to the use of the system for matrix-matrix multiplication, the solution of linear matrix-matrix equations and matrix inversion. We then consider in Section 8 its use in the solution of nonlinear matrix equations with specific application to the linear-quadratic-regulator problem and algebraic Riccati equation of optimal control engineering.

2. ITERATIVE OPTICAL PROCESSOR SYSTEM

The general schematic of the iterative optical processor (IOP) is shown in Figure 1. This figure illustrates the use of the IOP both as a vector-matrix multiplier and as an iterative processor for the solution of linear vector-matrix equations. Since an understanding of this architecture is paramount to the remainder of this paper, we review its operations. To multiply a vector by a matrix, the elements of the vector are used to spatially modulate a linear array of light emitting diodes (LEDs) or laser diodes (LDs) at P_1. We denote the light distribution leaving P_1 by the vector $\underline{x}$. Plane P_1 is imaged vertically onto P_2 and the output from each LED is expanded to uniformly illuminate the corresponding row of a 2-D mask at P_2. The light distribution leaving each column of P_2 is then summed onto a different photodetector at P_3. With the transmittance of P_2 specified by the matrix $\underline{H}$, the vector output from the linear photodetector array at P_3 is the matrix-vector product $\underline{H}\,\underline{x}$ as described in [1] and earlier by others [7,8].

In our system [2,3], we have included feedback of the photodetector outputs to the LED inputs through an electronic feedback system. In our original description of this system [2,3], the mask was $[\underline{I} - \underline{H}]$, where $\underline{I}$ is the identity matrix. The P_3 output is then $[\underline{I} - \underline{H}]\underline{x}(j)$, where $\underline{x}(j)$ denotes the vector $\underline{x}$ at the j-th iteration. An external vector $\underline{y}$ was added to this matrix-vector product in the electronic feedback system to form the new iterative input $\underline{x}(j + 1)$. Our system thus implemented the iterative algorithm

$$\underline{x}(j + 1) = [\underline{I} - \underline{H}]\underline{x}(j) + \underline{y} = \underline{x}(j) - \underline{H}\,\underline{x}(j) + \underline{y}. \tag{1}$$

In the steady-state when $\underline{x}(j + 1) = \underline{x}(j)$, equation (1) reduces to $\underline{H}\,\underline{x} = \underline{y}$ and the output $\underline{x}$ is the solution

$$\underline{x} = \underline{H}^{-1}\underline{y} \tag{2}$$

of the vector-matrix equation

$$\underline{H}\,\underline{x} = \underline{y}. \tag{3}$$

In the newest [5,6] version of this IOP, we have: (1) used fiber optics to realize the required projection from P_1 onto P_2 (this greatly improves the system's accuracy as well as its mechanical and positional stability and reduces its size and weight); (2) placed the photodetector at P_3 in direct contact with the matrix mask at P_2 (this is possible when the 4 mm height of each photodetector is matched to the height of the matrix mask at P_2. This further reduces the size and weight of the system and makes it even more stable for airborne applications); (3) included a microprocessor system with an arithmetic logic unit (ALU), memory and hardwired multiplier in the electronic feedback loop (Figure 2), (this increases the system's flexibility and versatility); and (4) modified the feedback system and the iterative algorithm to incorporate an acceleration parameter to insure convergence of the iterative algorithm (Section 3).

When funding permits, we plan: (1) to increase the size of the P_1 input and the P_2 mask from its present 10 × 10 level; (2) to incorporate a real-time and reuseable spatial-light modulator electro-optical mask element (such as a CCD-addressed liquid crystal light valve [9]) into the system; and (3) to increase the repertoire of operations and applications for the system. In Sections 4-8, we advance the first description of a selected number of general operations and problems as well as applications for which this new optical processor can be used.

3. CONVERGENCE AND HANDLING BIPOLAR AND COMPLEX-VALUED DATA

In this section, we first detail how this non-coherent optical system can be used to operate on bipolar and complex-valued data [5,6]. Since the LED outputs at P_1, the transmittance of the mask elements at P_2 and the photodetector outputs at P_3 are all real and positive, and since the dynamic range of the mask at P_2 is finite, various scaling, biasing and data encoding techniques are required to process vectors and matrices with bipolar and complex-valued data on the system. To handle complex-valued data, we partition the matrix $\underline{H}$ into its real and imaginary parts $\underline{H}_r$ and $\underline{H}_i$, respectively, as

$$\underline{H} = \begin{bmatrix} \underline{H}_r & -\underline{H}_i \\ \underline{H}_i & \underline{H}_r \end{bmatrix} , \tag{4}$$

where $\underline{H}_r$ and $\underline{H}_i$ are bipolar. This requires a P_2 mask with four times the space-bandwidth product of $\underline{H}$, an input LED array with twice the number of elements in $\underline{x}$ and a linear detector array with twice the number of elements present in $\underline{y}$.

To handle bipolar data, we decompose the input vector $\underline{x}$ into its positive and negative components. Let $\underline{a}^+$ and $\underline{a}^-$ denote the positive and negative components of the input vector $\underline{a}$ used in the actual system. The M elements a_{1m} and a_{2m} of $\underline{a}^+$ and $\underline{a}^-$ are generated according to

$$\left.\begin{aligned} a_{1m} &= 0.5\left(x_m + |x_m|\right) \\ a_{2m} &= 0.5\left(x_m - |x_m|\right) \end{aligned}\right\} , \tag{5}$$

where the x_m are the elements of the bipolar physical vector $\underline{x}$.

The system is then operated twice, first with $\underline{a}_1$ as the input and then with $\underline{a}_2$ as the input (with the same physical mask $\underline{B}$ used for all cycles). The outputs from the system on successive cycles are $\underline{B}\,\underline{a}_1$ and $\underline{B}\,\underline{a}_2$. These outputs are: (1) subtracted; (2) scaled by $(\bar{h} - \underline{h})$, where $\bar{h}$ and $\underline{h}$ are the maximum and minimum values of the elements of the matrix $\underline{H}$; and (3) biased by $\underline{h}\sum_{m=1}^{M} x_m$. The system's output $\underline{y}$ after two successive cycles is thus

$$\underline{y} = \underline{H}\,\underline{x} = \left(\bar{h} - \underline{h}\right)\left[\underline{B}\,\underline{a}_1 - \underline{B}\,\underline{a}_2\right] + \left(\underline{h}\sum_{m=1}^{M} x_m\right)[11\ldots1]^T. \tag{6}$$

To insure that the transmittance of the physical mask $\underline{B}$ has a transmittance for each element satisfying $0 \le b_{mn} \le 1$, we scale and bias $\underline{B}$ such that

$$b_{mn} = \frac{h_{mn} - \underline{h}}{\bar{h} - \underline{h}} , \tag{7}$$

where h_{mn} denotes the values of the elements of $\underline{H}$ and b_{mn} denotes the transmittances of the physical mask $\underline{B}$ placed at P_2.

The second issue to be detailed in this section is how rapid convergence of the iterative algorithm is insured. This is achieved by modifying our original iterative algorithm to include an acceleration parameter ω (as shown in Figure 1). The new

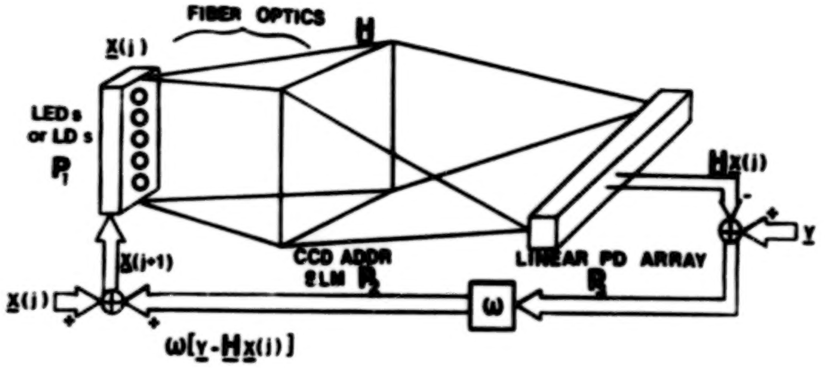

FIBER OPTICS
H
x(i)
LEDs or LDs
P1
Hx(i)
-
+
y
CCD ADDR SLM P2
LINEAR PD ARRAY P3
x(i+1)
x(i)
+
+
ω
ω[y-Hx(i)]

FIGURE 1 Schematic diagram of the iterative optical processor.

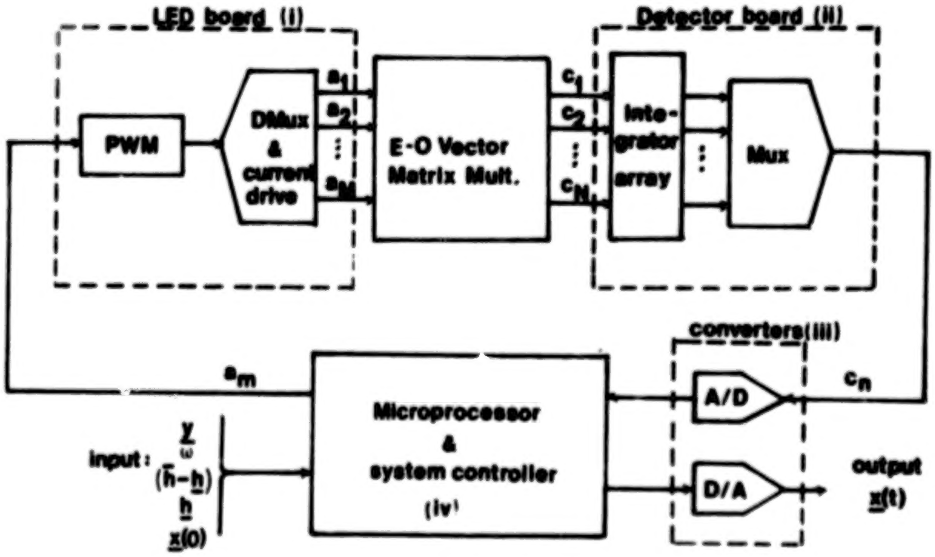

LED board (i)
Detector board (ii)
PWM
DMux & current drive
a1
a2
aM
E-O Vector Matrix Mult.
c1
c2
cN
Integrator array
Mux
converters(iii)
am
cn
A/D
D/A
Microprocessor & system controller (iv)
input: y/ω (h̄-h) h x(0)
output x(t)

FIGURE 2 Schematic diagram of the microprocessor-based fiber-optic iterative optical processor.

system thus implements the iterative algorithm [4-6]

$$\underline{x}(j+1) = \underline{x}(j) + \omega[\underline{y} - \underline{H}\,\underline{x}(j)]. \tag{8}$$

We select ω by one of the following criteria to insure rapid convergence of the algorithm. We can select

$$\omega = -1/\lambda_{max}, \tag{9}$$

where λ_{max} is the maximum eigenvalue, in absolute value, of the matrix $\underline{H}$, or we can obtain an approximation to (9) by [10]

$$\lambda_{max} \leq ||\underline{H}|| = \left[\sum_{m=1}^{M} \sum_{n=1}^{N} h_{mn}^2\right]^{1/2}, \tag{10}$$

where $||\underline{H}||$ is the Euclidean norm of the ($N \times N$) matrix $\underline{H}$. The ω criterion resulting from the upper-bound in (10) can be modified by selecting $\omega = -K/\lambda_{max}$, where $K > 1$ is a constant selected empirically from analysis of a specific problem. (In several specific case studies that we have considered, $K \approx 2$-3 was found to be adequate.) The microprocessor feedback system (Figure 2) performs the necessary scaling, biasing and preprocessing described by (5) and (7), the detector post-processing in (6), and the acceleration parameter selection noted in (10). In Section 6, we describe how the IOP itself can be used to calculate the acceleration parameter ω in (9).

4. SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS

As our first general IOP application, we consider the use of the system of Figure 1 for the solution of simultaneous linear (difference, differential or algebraic) equations. The general iterative algorithm for the solution of linear difference equations is

$$\underline{x}(j+1) = \underline{\Phi}\,\underline{x}(j) + \underline{f}, \tag{11}$$

where j is the discrete-time index, $\underline{f}$ is the forcing function equal to a constant vector (or, more generally, a vector of time-functions), $\underline{x}(j)$ is the state of the system, and $\underline{\Phi}$ is the open-loop system matrix.

The first application of the IOP of Figure 1 to the problem in (11) is as a dynamic system simulator. In this case, we model the physical system by the state-space differential equation model

$$\frac{d\underline{x}}{dt} = \underline{A}\,\underline{x} + \underline{b}, \tag{12}$$

where $\underline{A}$ is the open-loop system matrix and the vector $\underline{b}$ is a constant or function of time. We then utilize a numerical analysis algorithm to discretize this system model. We illustrate the approach by the forward-Euler approximation [11]

$$\left.\frac{d\underline{x}(t)}{dt}\right|_{t=jT} \simeq \frac{\underline{x}(j+1) - \underline{x}(j)}{T}\,, \tag{13}$$

where T is the constant time-increment (or step-size) between discrete samples described by the index j. Substituting (13) into (12), we can simulate the physical system modeled by (12) on the IOP of Figure 1 by the linear iterative algorithm

$$\underline{x}(j+1) = [\underline{I} + T\,\underline{A}]\,\underline{x}(j) + T\,\underline{b}\,. \tag{14}$$

By analogy with (11), $[\underline{I} + T\,\underline{A}] = \underline{\Phi}$ and $T\,\underline{b} = \underline{f}$. Other numerical analysis algorithms, such as the trapezoidal rule, are possible. The most preferable ones appear to be the Runge-Kutta and predictor-corrector algorithms [12].

The second application of (11) on the IOP is the iterative solution of linear algebraic equations. In this case, the iterative system algorithm of (8) is used. When rearranged in the form in (1), we can identify the $\underline{\Phi}$ matrix and the $\underline{f}$ vector in (11) as $[\underline{I} - \omega\underline{H}] = \underline{\Phi}$ and $\omega\underline{y} = \underline{f}$. The algorithms in (8) and (1) have been used [2,13] to calculate the set of adaptive weights necessary in adaptive phased-array radar signal processing. While space does not allow us to elaborate on this application, we note that, in this case, the matrix $\underline{H}$ in (3) is the covariance matrix $\underline{M}$ of the far field input to the antenna, $\underline{y}$ is the steering vector $\underline{s}$ for the antenna and the unknown vector $\underline{x}$ to be determined is the set of adaptive weights $\underline{w}$ to be applied to the received signals to steer the antenna in the desired direction (defined by $\underline{s}$) and to null the noise field (defined in frequency, velocity, angle and range by $\underline{M}$).

Four iterative algorithms to solve (3) for $\underline{x}$ given by (2) can directly be identified. The first is the Richardson algorithm [14] of (1) implemented with the acceleration parameter ω as in (8). The remaining three algorithms can be described in a new and quite useful formulation by decomposing the matrix $\underline{H}$ into the sum of a diagonal matrix $\underline{D}$ and lower and upper triangular matrices $\underline{L}$ and $\underline{U}$ as

$$\underline{H} = \underline{D} - \underline{L} - \underline{U}\,. \tag{15}$$

In terms of the $\underline{D}$, $\underline{L}$ and $\underline{U}$, we can write the remaining three algorithms as [15]:

$$\text{Jacobi:}\quad \underline{x}(j+1) = \left[\underline{D}^{-1}(\underline{L} + \underline{U})\right]\,\underline{x}(j) + \underline{D}^{-1}\underline{y} \tag{16a}$$

$$\text{Gauss-Seidel:}\quad \underline{x}(j+1) = \left[(\underline{D} - \underline{L})^{-1}\underline{U}\right]\underline{x}(j) + (\underline{D} - \underline{L})^{-1}\,\underline{y} \tag{16b}$$

$$\text{Overrelaxation:}\quad \underline{x}(j+1) = \left\{(\underline{D} - \omega\underline{L})^{-1}[(1-\omega)\underline{D} + \omega\underline{U}]\right\}\underline{x}(j) + \omega(\underline{D} - \omega\underline{L})^{-1}\underline{y}. \tag{16c}$$

The choice of one of the four algorithms in (8) or (16) depends on many factors that are highly application and problem dependent (e.g., convergence of the algorithm, dy-

dynamic range of the mask, number of iterations required and the need for an adaptive mask).

5. LEAST-SQUARES PROBLEMS

The solution to many physical and statistical problems can be formulated in the general context of a least mean-squares problem. A model is postulated to approximate measured data and the parameters of the model are selected to minimize the mean-square error between the measured data and the model. Curve-fitting and linear and polynomial regression are classical examples of least mean-squares problems. We formulate and address below the solution on our IOP of this class of important problems.

We begin by reformulating our simultaneous linear algebraic equation problem solution in Section 4 as a least-squares problem. In the context of Figure 3, the input to the operator $\underline{H}$ is $\underline{x}$ and noise or an error source $\underline{\varepsilon}$ is present at the output such that the observable output $\underline{y}$ differs from $\underline{H}\,\underline{x}$. The problem is thus to find the $\underline{x}$ that minimizes the square of the difference

$$J = ||\underline{y} - \underline{H}\,\underline{x}||^2 \tag{17}$$

between $\underline{y}$ and $\underline{H}\,\underline{x}$.

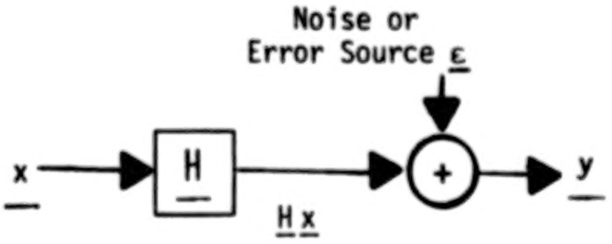


FIGURE 3 Schematic diagram of a least-mean squares problem formulation.

Upon expanding (17), we obtain

$$J = [\underline{y} - \underline{H}\,\underline{x}]^T [\underline{y} - \underline{H}\,\underline{x}] = \underline{y}^T\underline{y} - \underline{y}^T\underline{H}\,\underline{x} - \underline{x}^T\underline{H}^T\underline{y} + \underline{x}^T\underline{H}^T\underline{H}\,\underline{x}\,. \tag{18}$$

To minimize (18), we set the partial derivative or gradient of the scalar performance index in (17) with respect to the unknown vector $\underline{x}$ equal to zero. The resultant column vector is then the solution of

$$\left.\begin{array}{c} -2\underline{H}^T\underline{y} + 2\underline{H}^T\underline{H}\,\underline{x} = \underline{0} \\ \text{or} \\ \left[\underline{H}^T\underline{H}\right]\underline{x} = \underline{H}^T\underline{y} \end{array}\right\} \quad . \tag{19}$$

In least-squares data processing, $\underline{y}$ is an M-vector, $\underline{H}$ is an $(M \times N)$ matrix and $\underline{x}$ is an N-vector, where $M \geq N$. We consider first the case when the number of unknowns N equals the number of equations M. In this case, $\underline{H}$ is a square matrix and the solution to (19) becomes

$$\underline{x} = \left[\underline{H}^T\underline{H}\right]^{-1} \underline{H}^T \underline{y} = \underline{H}^{-1}\underline{H}^{-T}\underline{H}^T\underline{y} = \underline{H}^{-1}\underline{y} \ , \tag{20}$$

where $\underline{H}$ is assumed by implication to be a non-singular matrix and therefore to be invertible. This occurs when the $\underline{y} = \underline{H}\,\underline{x}$ problem being solved is well-formulated as occurs in all practical engineering problems. Thus, in this case, the solution $\underline{x} = \underline{H}^{-1}\underline{y}$ minimizes the quadratic performance index in (17). For this case, we solve the least mean-square problem by our Richardson algorithm of (8).

A more interesting problem (corresponding to the practical case of curve-fitting) occurs when there are more equations M than unknowns N (i.e., when $M > N$). In this case, $\underline{H}$ is non-square and therefore non-invertible. We thus solve (19) by the new iterative algorithm

$$\underline{x}(j+1) = \underline{x}(j) + \omega\left\{\left[\underline{H}^T\underline{y}\right] - \left[\underline{H}^T\underline{H}\right]\underline{x}(j)\right\} \ . \tag{21}$$

This is equivalent to pre-multiplying (3) by $\underline{H}^T$ and applying our iterative Richardson algorithm in (8) to the solution of $[\underline{H}^T\underline{H}]\underline{x} = \underline{H}^T\underline{y}$ rather than $\underline{H}\,\underline{x} = \underline{y}$.

The least-squares fitting of experimental observations leads to a more general iterative vector-matrix solution. Suppose that we have L data points $\underline{p} = \{p_\ell\}$ and wish to approximate the set of observations $z(\underline{p})$, in the least-squares sense, by the finite linear combination

$$z(\underline{p}) \simeq \underline{c}^T\underline{\phi}(\underline{p}) \tag{22}$$

of basis functions $\underline{\phi} = \{\phi_n\}$ with the associated weighting coefficients $\underline{c} = \{c_n\}$. In (22), each of the N basis functions $\underline{\phi}$ are evaluated at the L data points $\{p_\ell\}$ and the relationship in (22) is approximate if there are more data points L than basis functions N. The errors or residuals $[z(\underline{p}) - \underline{c}^T\underline{\phi}(\underline{p})]$ of the curve-fit are the differences between the observed data $z(\underline{p})$ and the approximation $\underline{c}^T\underline{\phi}(\underline{p})$. According to the principle of least-squares, we select the weighting coefficients in (22) to minimize the sum-of-squares of the residuals. To find the coefficients $\underline{c}$, therefore, we minimize the mean-square difference

$$J = \sum_{\substack{\text{Data} \\ \{p_\ell\}}} \left[z(p_\ell) - \underline{c}^T\underline{\phi}(p_\ell)\right]^2 \ . \tag{23}$$

We proceed as before. Upon setting $\partial J/\partial \underline{c} = \underline{0}$, we find

$$\underbrace{\left[\sum_{\substack{\text{Data}\\ \{p_\ell\}}} \underline{\phi}(p_\ell)\underline{\phi}^T(p_\ell\right]}_{\underline{H}} \underbrace{\underline{c}}_{\underline{x}} = \underbrace{\left[\sum_{\substack{\text{Data}\\ \{p_\ell\}}} \underline{\phi}(p_\ell)z(p_\ell)\right]}_{\underline{y}} \tag{24}$$

which is again of the general form $\underline{H}\,\underline{x} = \underline{y}$, with the summation over the $\{p_\ell\}$ data points incorporated into the matrix and vectors in the algorithm in (21). When the basis functions $\underline{\phi} = \{\phi_n\}$ are specified, the matrix $\underline{H}$ can be precomputed and the problem solved by the iterative algorithm in (8) or (21). A simple discrete example of $\{\phi_n\}$ arises in the calculation of the best straight-line fit through a set of data points. In this case, $\phi_1 = 1$ and $\phi_2 = p$ and $z(p) = c_1 + c_2 p$. A simple continuous example case is the Fourier series expansion of $z(p)$, in which case the $\{\phi_n\}$ are the complex exponentials $\exp(\pm jn\omega_0 p)$.

In this section we have illustrated and formulated the least mean-squares problem as the iterative solution of the system of linear algebraic equations $\underline{H}\,\underline{x} = \underline{y}$ on the IOP. We recognize that the range of applications of our formulation includes curve-fitting, linear and nonlinear regression, state and parameter estimation, orbit determiniation and signal processing in control and communication engineering. We will address these applications in our future work.

6. DECONVOLUTION AND EIGENVALUE PROBLEMS

The need to implement a deconvolution frequently arises. In such applications (Figure 4), the measured output data $\underline{y} = \{y_m\}$ from a system characterized by the impulse response $\{h_m\}$ will be a modified version of the original input vector $\underline{x} = \{x_n\}$; i.e.,

$$y_m = \sum_{n=0}^{m} h_{(m-n)}\, x_n \,. \tag{25}$$

In (25), we assume that the system is causal, otherwise we appropriately shift the impulse response so that $h_m = 0$ for $m < 0$. The solution of (25) for the unknown input $\underline{x}$ is directly described by the vector-matrix equation $\underline{y} = \underline{H}\,\underline{x}$, where (for the case of $M = N = 3$) the matrix $\underline{H}$ is shown below:

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 \\ h_1 & h_0 & 0 \\ h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \end{bmatrix} . \tag{26}$$

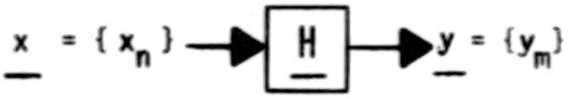


FIGURE 4 Simplified schematic diagram of a deconvolution processing problem.

In the general case of $M > N = 3$, the size of $\underline{H}$ grows, but $\underline{H}$ still has only three non-zero diagonals (with equal values h_0, h_1 and h_2, respectively, along each diagonal), and there are an additional $(M - N)$ elements with zero value added to $\underline{x}$. The solution of the deconvolution problem in (25) is thus directly realizable on the IOP by the iterative algorithm in (8) with the matrix mask in the form shown in (26). Once again, we see how different problems can be solved on the same IOP by different choices of the matrix mask and the iterative algorithm used.

Another quite general and useful matrix operation is the calculation of the eigenvalues λ_n and corresponding eigenvectors $\underline{\phi}_n$ of the $(N \times N)$ matrix $\underline{H}$. This operation can be performed by the iterative algorithm [16]

$$\underline{x}(j + 1) = \underline{H}\,\underline{x}(j) \ , \tag{27}$$

which is equivalent to (8) with the exogenous vector $\underline{y}$ set equal to zero. To see how (27) allows calculation of the λ_n and $\underline{\phi}_n$, we assume that $|\lambda_1|$ is the largest eigenvalue and that the eigenvalues are ordered, in absolute value, according to $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_N|$. By singular value decomposition, we write $\underline{H}$ and the original initialization vector $\underline{x}_1(0)$ as

$$\left.\begin{aligned} \underline{H} &= \sum_{i=1}^{N} \underline{\phi}_i \lambda_i \underline{\phi}_i^T \\ \text{and} \qquad & \\ \underline{x}_1(0) &= \sum_{\ell=1}^{N} a_\ell \underline{\phi}_\ell \end{aligned}\right\} . \tag{28}$$

After j iterations, $\underline{x}(j + 1) = \underline{H}^j \underline{x}_1(0)$ is obtained. Substituting (28) into this expression, we find (for large j)

$$\underline{x}(j + 1) = \underline{H}^j \underline{x}_1(0) \underset{\text{Large } j}{\approx} a_1 \lambda_1^j \underline{\phi}_1 \ . \tag{29}$$

We form the component-wise ratio $x_i(j + 1)/x_i(j)$ and from it find λ_1 (the largest eigenvalue of $\underline{H}$). We divide the denominator of the output vector $\underline{x}$ by the sum of the squares of its elements and thus obtain the principal eigenvector $\underline{\phi}_1$. We then use the new initial vector $\underline{x}_2(0) = \underline{x}_1(0) - a_1\underline{\phi}_1$. Repeating the same iterative procedure in

(27), we then find λ_2 and $\underline{\phi}_2$. By continuing this process, all of the eigenvalues and eigenvectors of $\underline{H}$ can be determined in decreasing order. Modifications to this procedure allow us to find the eigenvector whose eigenvalue lies closest to a prescribed value [16]. Many other extensions of these techniques and applications of the above results are possible and will be the subject of future work.

7. SOLUTION OF MATRIX EQUATIONS

Many problems involve matrix-matrix multiplication. There are two ways to realize a matrix-matrix multiplication on our vector-matrix multiplier. In the case of large matrices, the preferable technique is to form the product $\underline{Y} = \underline{M} \cdot \underline{N}$ of the matrices $\underline{M}$ and $\underline{N}$ by feeding sequentially the columns $\underline{n}_n$ of $\underline{N}$ as successive input vectors. We thus realize the matrix-matrix product by performing N vector-matrix multiplications; i.e.,

$$\underline{Y} = \underline{M} \cdot \underline{N} = \underline{M}[\underline{n}_1\, \underline{n}_2 \,.\,.\,.\, \underline{n}_N] \, , \tag{30}$$

where the matrix output $\underline{Y}$ appears sequentially as N column vectors; i.e., $\underline{Y} = [\underline{y}_1\, \underline{y}_2 \,.\,.\,.\, \underline{y}_N]$. A second technique is to vectorize the matrix $\underline{N}$ into an N^2 element vector whose first N elements are the elements of $\underline{n}_1$ and the next N elements are the elements of $\underline{n}_2, \ldots$. The matrix $\underline{M}$ is formatted as an $(N^2 \times N^2)$ matrix whose diagonal blocks are replications of the matrix $\underline{M}$; i.e.,

$$\underline{M} \cdot \underline{N} = \begin{bmatrix} \underline{M} & & \bigcirc \\ & \underline{M} & \\ & & \ddots & \\ \bigcirc & & & \underline{M} \end{bmatrix} \begin{bmatrix} \underline{n}_1 \\ \underline{n}_2 \\ \vdots \\ \underline{n}_N \end{bmatrix} = \begin{bmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_N \end{bmatrix} \, . \tag{31}$$

With a matrix-matrix multiplier realized by (30) or (31), we can thus use the IOP to solve matrix-matrix equations such as

$$\underline{H}\,\underline{X} = \underline{Y} \tag{32}$$

by vectorizing the matrix $\underline{Y}$ or by operating the system on sequential cycles with the column vectors of $\underline{Y}$ as successive inputs. A particularly useful operation that is now possible is matrix inversion. In this case, we solve (32) using (8), but with $\underline{Y} = \underline{I}$ (the identity matrix). A final useful matrix equation that frequently arises is the solution for the embedded matrix $\underline{X}$ in an equation of the form

$$\underline{A}\,\underline{X}\,\underline{B} = \underline{Y} \, . \tag{33}$$

In this case, we write the matrices $\underline{Y}$ and $\underline{X}$ as the column vectors $\underline{C}[\underline{Y}]$ and $\underline{C}[\underline{X}]$, whose elements are the **lexigraphically-ordered** elements of the matrices $\underline{Y}$ and $\underline{X}$, respectively. We solve (33) for $\underline{X}$ by writing (33) in the form

$$[\underline{A} \otimes \underline{B}^T]\underline{C}[\underline{X}] = \underline{C}[\underline{Y}] \, , \tag{34}$$

where $\otimes$ denotes the outer or Kronecker product. For the case of the (2×2) matrix $\underline{A}$,

$$\underline{A} \otimes \underline{B}^T = \begin{bmatrix} a_{11}\,\underline{B}^T & a_{12}\,\underline{B}^T \\ a_{21}\,\underline{B}^T & a_{22}\,\underline{B}^T \end{bmatrix} . \tag{35}$$

Solution of the Lyapunov matrix equation $\underline{X}\,\underline{A} + \underline{A}^T\,\underline{X} = \underline{Y}$ is now possible on the vector-matrix IOP.

8. SOLUTION OF NONLINEAR MATRIX EQUATIONS

In Section 7, we described how we have broadened the repertoire of operations achievable on the IOP to include the full-class of linear and embedded matrix equations. As our final general iterative vector matrix operation, we consider its use in the solution of nonlinear matrix equations. The need for the solution of such equations arose in our original use of the system for adaptive phased-array radar processing [2] and in our present optimal control linear-quadratic-regulator application [4].

The general problem of two quadratic equations in the two unknowns p and q can be written as

$$f_n(p,q) = a_n p^2 + b_n pq + c_n q^2 + d_n p + e_n q + r_n = 0 \tag{36}$$

for $n = 1$ and 2. We rewrite this pair of two nonlinear equations in vector form as

$$\underline{f}[\underline{x}] = \underline{0} \tag{37}$$

where $\underline{f}[\underline{x}] = [f_1(\underline{x})\ \ f_2(\underline{x})]^T$ and $\underline{x} = [p\ \ q]^T$. We then solve (37) by the Newton-Raphson iterative algorithm

$$\underline{x}(j+1) = \underline{x}(j) - \left[\frac{\partial \underline{f}}{\partial \underline{x}}\right]^{-1}_{\underline{x}(j)} \underline{f}[\underline{x}(j)]. \tag{38}$$

The solution of (38) requires two iterative loops. The Jacobian matrix $\underline{J} = [\partial \underline{f}/\partial \underline{x}]$ is stored algebraically. At each iteration, $\underline{J}$ is evaluated numerically with dedicated electronics at the last iterate $\underline{x}(j)$. The inverse matrix $[\underline{J}[\underline{x}(j)]]^{-1}$ is evaluated on the optical vector-matrix system in an inner iterative loop. The new $\underline{x}(j+1)$ iterate is then evaluated optically in an outer iterative loop.

Three immediate applications for such a nested two-loop iterative algorithm arise. The first occurs in the implementation of the overrelaxation linear algebraic equation solution of (16c) with $[\underline{D} - \omega\underline{L}]$ on the left-hand side of the equation. In this case, a matrix-vector multiplication is required to obtain the external vector to be added. The second case arises in the use of (8) to solve for the adaptive weights $\underline{w}$ for a phased-array radar whose noise field is described by the covariance matrix $\underline{M}$ and the direction of its main lobe is defined by the steering vector $\underline{s}$. In this case, we solve the vector-matrix equation $\underline{M} \cdot \underline{w} = \underline{s}$. When the noise distribution varies, the matrix mask $\underline{M}$ must be updated.

The IOP application we are presently considering under support from the NASA-Lewis Research Center is the use of the IOP in the solution of the linear-quadratic-regulator (LQR) problem of optimal control. In this application, the IOP is used to calculate the optimal controls to be applied to an F100 aircraft engine. To determine these control signals, we must solve the algebraic Riccati equation and calculate the LQR feedback gains. In this application, we use: (1) the vectorization of a matrix; (2) the Kronecker product technique; and (3) the nested inner and outer loop system to solve the nonlinear algebraic Riccati equation [4].

9. SUMMARY

In this paper, the general-purpose nature of an iterative vector-matrix processor has been emphasized. This system is capable of solving a wealth of general purpose applications. General operations described included: linear difference and differential equations, linear algebraic equations, matrix equations, matrix inversion, nonlinear matrix equations, deconvolution and eigenvalue and eigenvector computations. Engineering applications being addressed for these different operations and for the IOP are: adaptive phased-array radar, time-dependent system modeling, deconvolution and optimal control.

ACKNOWLEDGEMENTS

The authors thank Dr. Mark Carlotto for fabrication of the laboratory IOP and for the demonstration of its use in many of the applications described in this paper. We also thank the Air Force Office of Scientific Research for intermediate support of the general applications of this architecture and the NASA-Lewis Research Center for present support of this data processing concept for linear-quadratic-regulator optimal control applications.

REFERENCES

1. J. Goodman et al, *Optics Letters*, *2*, 1 (1978).

2. D. Psaltis et al, *SPIE*, *180*, 114 (1979).

3. D. Psaltis et al, *Optics Letters*, *4*, 348 (1979).

4. D. Casasent et al, *SPIE*, *295*, (1981).

5. M. Carlotto and D. Casasent, A Fiber Optic Microprocessor-Based Vector Matrix Optical Processor. *Applied Optics*, [1982].

6. M. Carlotto and D. Casasent, An Iterative Optical Vector-Matrix Processor. *Transformations in Optical Signal Processing*, W.T. Rhodes, J.R. Fienup, and B.E.A. Saleh, editors, Soc. of Photo-Optical Instrumentation Engineers, Dellingham, Washington, [1982].

7. P. Mengert et al, U.S. Patent #3,525,856, filed Oct, 6, 1966.

8. A. Edison and M. Noble, *Optical Analog Matrix Processors*, AD646060 (Nov. 1966).

9. J. Grinberg et al, *Proc. Soc. Photo. Instr. Engrs.*, *128*, 253 (1977).

10. E. Kreyszig, *Advanced Engineering Mathematics*, John Wiley & Sons, Inc., New York, 1972, p. 686.

11. L.O. Chua and P.M. Lin, *Computer-Aided Analysis of Electronic Circuits*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975, Chapter 1.

12. L.O. Chua and P.M. Lin, *Computer-Aided Analysis of Electronic Circuits*, Prentice-Hall, Englewood Cliffs, New Jersey, 1975, Chapter 11.

13. D. Casasent, Optical Processing for Adaptive Phased-Array Radar - Final Technical Report. RADC-TR-81-152, June 1981.

14. L. Richardson, *Philos. Trans. Roy. Soc. London, Ser A210*, 307 (1910).

15. R.S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962, pp. 56-59.

16. B. Kumar and D. Casasent, Eigenvector Determination by Iterative Optical Methods. *Applied Optics*, [Nov. 1981].

INTERFEROMETRIC SURFACE-WAVE ACOUSTO-OPTIC TIME-INTEGRATING CORRELATORS

Norman J. Berg, Irwin J. Abramovitz, and Michael W. Casseday
U.S. Army Electronics Research and Development Command
Harry Diamond Laboratories, 2800 Powder Mill Road
Adelphi, Maryland 20783

ABSTRACT

A novel structure for a coherent-interferometric acousto-optic (AO) time-integrating correlator has been implemented by using a single surface acoustic wave (SAW) device with tilted transducers to reduce intermodulation terms. The SAW device was fabricated on Y-Z $LiNbO_3$ with a center frequency of 175 MHz, a bandwidth of 60 MHz, and a time aperture of about 10 μs. The instantaneous bandwidth of the AO correlator is determined by the spacing density of the photodetector array, with a potential of 120 MHz. Typical integration times are 30 to 40 ms, providing processing gains in excess of 10^6. Such a device is very useful in providing fast synchronization of communication links. In addition, the device can demodulate to base band and simultaneously act as a synchronization lock monitor for moderate data rates. Where processing may be limited by doppler shifts, a two-dimensional architecture has been implemented to allow full processing gain.

INTRODUCTION

Current digital and microwave technology has made possible wideband communications for space applications. These systems present unique problems for which acousto-optics (AO) may provide solutions. The relative ease in applying multiple transducers to surface acoustic wave (SAW) delay lines allows novel architectures for such signal processing functions as correlation or convolution. Where large processing gain is required, integration in time rather than space permits time-bandwidth products in excess of 10^6. Coherent, interferometric schemes provide both time (e.g., time-difference-of-arrival) and frequency information simultaneously. Two one-dimensional, SAW AO time-integrating correlators and a two-dimensional correlator incorporating these features have been constructed, and the results are presented.

ONE-DIMENSIONAL CORRELATOR ANALYSIS AND APPLICATIONS

Two-input-beam time-integrating correlator

Figure 1 illustrates the operation of a two-beam SAW AO correlator.[1] A transducer is deposited on each end of the delay line (y-cut, z-propagating lithium niobate) at an angle with respect to the perpendicular to the z-axis of Θ_{Bn}, the Bragg angle in the delay line material[2] at the correlator design center frequency, ω_0. The relative tilt between the two transducers is $2\Theta_{Bn}$. The correlator input signals, $A(t) \cos \omega_A t$ and $B(t) \cos \omega_B t$, generate counter-propagating SAW's which interact with two sheet beams of laser light generated from a single laser and projected into the top side of the delay line with an angle of $4\Theta_B$ between them. Here, Θ_B is the Bragg angle in air for the frequency ω_0.[2] Due to the strong angular dependence of the AO Bragg interation, the right sheet beam interacts primarily with the SAW launched by the right transducer; likewise, the left sheet beam interacts primarily with the SAW generated by the left transducer. Cross terms are down by 40 dB.[3] The diffracted light is imaged onto an integrating, square-law photodiode array. For equal sheet beams of uniform intensity the diffracted light may be described by

$$L_1(t,z) = A(t - \frac{z}{\upsilon}) \cos[\omega_\ell(t - \frac{z \sin 2\Theta_B}{c}) - \omega_A(t - \frac{z}{\upsilon})] \tag{1}$$

and

$$L_2(t,z) = B(t + \frac{z}{\upsilon}) \cos[\omega_\ell(t + \frac{z \sin 2\Theta_B}{c}) - \omega_B(t + \frac{z}{\upsilon})] . \tag{2}$$

Here, ω_ℓ is the light frequency, t is time, z is the distance along the delay line and along the photodiode array (center is $z = 0$), υ is the acoustic propagation velocity, and c is the free-space light velocity. The photodiode array output current is proportional to the square of the sum of $L_1(t,z)$ and $L_2(t,z)$. The significant term, the cross-product proportional to $L_1(t,z) \cdot L_2(t,z)$, may be manipulated trigonometrically to generate frequency sum and difference terms. Upon time integration, only the difference term remains, yielding an output voltage proportional to

$$V(t,z) = \int_T A(t - \frac{z}{\upsilon})\, B(t + \frac{z}{\upsilon}) \cos[\frac{2\omega_\ell z \sin 2\Theta_B}{c} + (\omega_A - \omega_B)t - (\omega_A + \omega_B)\frac{z}{\upsilon}]dt \quad . \tag{3}$$

Since $\sin 2\theta_B = (\omega_0/\upsilon) / (\omega_\ell/c)$, and for $\omega_A = \omega_B$ this output voltage becomes:

$$V(t,z) = \cos[\frac{2z}{\upsilon}(\omega_0 - \omega_A)] \int_T A(t - \frac{z}{\upsilon}) B(t + \frac{z}{\upsilon})dt. \tag{4}$$

Thus, the output voltage provides the correlation of the input signals (modulation and carrier) about the correlator design center frequency. Figure 2 shows the autocorrelation of a bi-phase signal with 2.5 MHz bandwidth at 1 MHz above ω_0. The bias level, produced by the square terms proportional to the integral of $L_1(t,z)^2 + L_2(t,z)^2$, depicts the Gaussian intensity profile of the sheet beams. This bias level may be easily removed by high pass filtering or by a subtraction of one output from a subsequent output produced with one input phase-inverted. The position of the fringe pattern is due to the relative time delay between the signals. Instantaneous bandwidths of 60 MHz at 175 MHz have been achieved with integration times of 30 to 40 ms for time apertures of 10 μs.

Single-input-beam time-integrating correlator

Since the angle of the incoming beam with respect to the SAW wavefront (transducer tilt) equals that for the diffracted beam, an inverse architecture using one input laser beam is possible, as shown in Figure 3. The output voltage from the photodiode array can be shown to be identical to that for the two-beam correlator as described by equation (4). Although intermodulation terms (the interaction of the SAW from the left transducer appearing at the right output diffracted beam and vice-versa) are suppressed by less than the 40 dB achieved with the two-beam architecture, suppression is sufficient for most applications.

The large processing gains, linearity, and time delay capabilities of the SAW time-integrating correlators make them useful as advanced signal processors. These correlators may be employed merely to provide synchronization in a communications system as a discrete version of the "sliding correlator"[4] or may be used to demodulate the signal as well as to continually monitor synchronization. Figure 4 shows the output of the correlator, with bias subtracted, for inputs consisting of an information modulated, hybrid bi-phase wideband signal and an unmodulated (information free) reference signal. The overall bandwidth was 16 MHz with a bi-phase rate of 1.3 million bits per second. Integration was sychronized with the information rate. Input signal-to-noise ratio was -30 dB.

Two-Dimensional Correlator

A two-dimensional, four product correlator has been demonstrated which features the interference between two doubly diffracted beams from perpendicular SAW delay lines. As shown in Figure 5, the initial source of the sheet beams is a single laser. Since the SAW devices are perpendicular, interaction in only one device is between light polarized perpendicular to the direction of travel of the SAW's in that device. This is the preferred polarization; for light polarized parallel to the direction of SAW propagation, the diffraction efficiency may be somewhat reduced and a broadening of the Bragg angle dependence is experienced.

The double diffraction effect[5] is diagrammed in Figure 6 for the horizontal SAW device. The light from the first diffraction is of the form:

$$L_{H_1}(t,z) = A(t - \frac{z}{\upsilon}) \cos[\omega_\ell(t + \frac{z \sin \phi_1}{c}) + \omega_A (t - \frac{z}{\upsilon})]. \tag{5}$$

The doubly diffracted light is then of the form:

$$L_{H_2}(t,z) = A(t - \frac{z}{\upsilon}) B(t + \frac{z}{\upsilon}) \cos[\omega_\ell(t + \frac{z \sin \phi_1}{c}) + \omega_A (t - \frac{z}{\upsilon}) + \omega_B(t + \frac{z}{\upsilon})]. \tag{6}$$

Since $\phi_1 = 2(\Theta_{B_1} - \Theta_{B_2})$, then $\sin \phi_1 = (\omega_1/\upsilon)/(\omega_\ell/c) - (\omega_2/\upsilon)/(\omega_\ell/c)$, and equation (6) reduces to:

$$L_{H_2}(t,z) = A(t - \frac{z}{\upsilon}) B(t + \frac{z}{\upsilon}) \cos[(\omega_\ell + \omega_A + \omega_B) t + \frac{z}{\upsilon} \Delta_H], \tag{7}$$

where $\Delta_H = (\{\omega_1 - \omega_2\} - \{\omega_A - \omega_B\})$. Here, ω_1 and ω_2 are design center frequencies for which the Bragg angles are Θ_{B_1} and Θ_{B_2}, respectively. ω_1 and ω_2 are selected to be different so that the doubly diffracted light may be separated from the undiffracted beam. Similarly, for the vertical device, the doubly diffracted light is

$$L_{V_2}(t,y) = C(t - \frac{y}{\upsilon}) D(t + \frac{y}{\upsilon}) \cos[(\omega_\ell + \omega_C + \omega_D) t + \frac{y}{\upsilon} \Delta_V], \tag{8}$$

where $\Delta_{V_2} = (\{\omega_3 - \omega_4\} - \{\omega_C - \omega_D\})$. Those doubly diffracted beams are imaged onto a photodiode area array or a vidicon that integrates the output current which is proportional to the square of the sum of these beams. The frequency difference term derived from the cross product, which is proportional to $L_{H_2}(t,z) \cdot L_{V_2}(t,y)$, produces an output voltage

$$V(t,z,y) = \int_T A(t - \tfrac{z}{\upsilon})\ B(t + \tfrac{z}{\upsilon})\ C(t - \tfrac{y}{\upsilon})\ D(t + \tfrac{y}{\upsilon})$$

$$\cos[(\omega_A + \omega_B - \omega_C - \omega_D)t + \tfrac{z}{\upsilon}\Delta_H - \tfrac{y}{\upsilon}\Delta_V)dt. \qquad (9)$$

If $(\omega_A + \omega_B) = (\omega_C + \omega_D)$, then equation (9) reduces to

$$V(t,z,y) = \cos[\tfrac{z}{\upsilon}\Delta_H - \tfrac{y}{\upsilon}\Delta_V] \int_T A(t - \tfrac{z}{\upsilon})$$

$$B(t + \tfrac{z}{\upsilon})\ C(t - \tfrac{y}{\upsilon})\ D(t - \tfrac{y}{\upsilon})\ dt. \qquad (10)$$

For $(\omega_1 - \omega_2) = (\omega_A - \omega_B)$ and $(\omega_3 - \omega_4) = (\omega_C - \omega_D)$, this further reduces to:

$$\int_T A(t - \tfrac{z}{\upsilon})\ B(t + \tfrac{z}{\upsilon})\ C(t - \tfrac{y}{\upsilon})$$

$$D(t + \tfrac{y}{\upsilon})dt. \qquad (11)$$

Thus, this device provides a two-dimensional, four-product correlation. It has the additional ability to measure some degree of deviation from design frequencies. The usefulness of this correlator for processing signals containing doppler shifts in frequency can be shown by replacing the generalized input signals to the vertical SAW delay line, $C(t) \cos \omega_C t$ and $D(t) \cos \omega_D t$, by linear FM chirps, $\cos(\omega_C + \alpha t)t$ and $\cos(\omega_D - \alpha t)t$. Equation (8) becomes:

$$L_{V_2}(t,y) = \cos[(\omega_\ell + \omega_C + \omega_D)t + \tfrac{y}{\upsilon}\Delta_H$$

$$- 4\alpha\,(\tfrac{y}{\upsilon})t]. \qquad (12)$$

The output voltage then becomes

$$V(t,z,y) = \int_T A(t - \tfrac{z}{\upsilon})\ B(t + \tfrac{z}{\upsilon}) \cos[\omega_A +$$

$$\omega_B - \omega_C - \omega_D)t + \tfrac{z}{\upsilon}\Delta_H - \qquad (13)$$

$$\tfrac{y}{\upsilon}\Delta_V + 4\alpha(\tfrac{y}{\upsilon})t]dt.$$

If $\omega_B = \omega'_B + \omega_{DP}$ where ω_{DP} is an unknown doppler shift and $\omega_A + \omega'_B = \omega_C + \omega_D$, then equation (13) reduces to

$$V(t,z,y) = \int_T A(t - \frac{z}{U})\ B(t + \frac{z}{U}) \cos[(\omega_{DP} + 4\alpha \frac{y}{U})t + \frac{z}{U} \Delta_H - \frac{y}{U} \Delta_V]dt. \qquad (14)$$

It can be seen that there is a y position for which $\omega_{DP} = -4\alpha \frac{y}{U}$ and the doppler shift is compensated. As an ambiguity function processor, correlation of a biphase code in the horizontal dimension provides range information, and doppler compensation in the vertical dimension provides velocity data. A four-product correlator has been implemented, and the output is shown in figure 7, where $_H$ and $_V$ are nonzero. The correlation spot has been moved from the center of the vidicon output by time delay (simulated range) and carrier frequency shift (simulated doppler). Bias subtraction has been performed using a video image processor.

SUMMARY

One- and two-dimensional interferometric time-integrating, acousto-optic correlators have been demonstrated. These offer large processing gains for application to wideband signal processing systems. Potential uses for synchronization, demodulation, and ambiguity function processing have been discussed.

References

1. N.J. Berg, I.J. Abramovitz, J.N. Lee, and M.W. Casseday, "A New Surface Wave Acousto-Optic Time Integrating Correlator," Appl. Phys. Lett., 36(4), pp. 256-258, 1980.

2. R. Adler, "Interaction Between Light and Sound," IEEE Spectrum 4, pp. 42-54, May 1967.

3. J.N. Lee, N.J. Berg, and M.W. Casseday, "Multichannel Signal Processing Using Acousto-optic Techniques," IEEE J. Quant. Electron ., QE-15 (11), pp. 1210-1215, Nov. 1979.

4. N.J. Berg, M.W. Casseday, I.J. Abramovitz, and J.N. Lee, "Radar and Communications Band Signal Processing Using Time-Integration Processors," SPIE, 232, Proc. 1980 IOCC, pp. [illegible]01-107, 1980.

5. J.N. Lee, N.J. Berg, and M.W. Casseday, "Multichannel Surface Acoustic Wave Correlation and Convolution with Acousto-Optic Processors," Proc. 1979 Ultrasonic Symposium, pp.34-39, 1979.

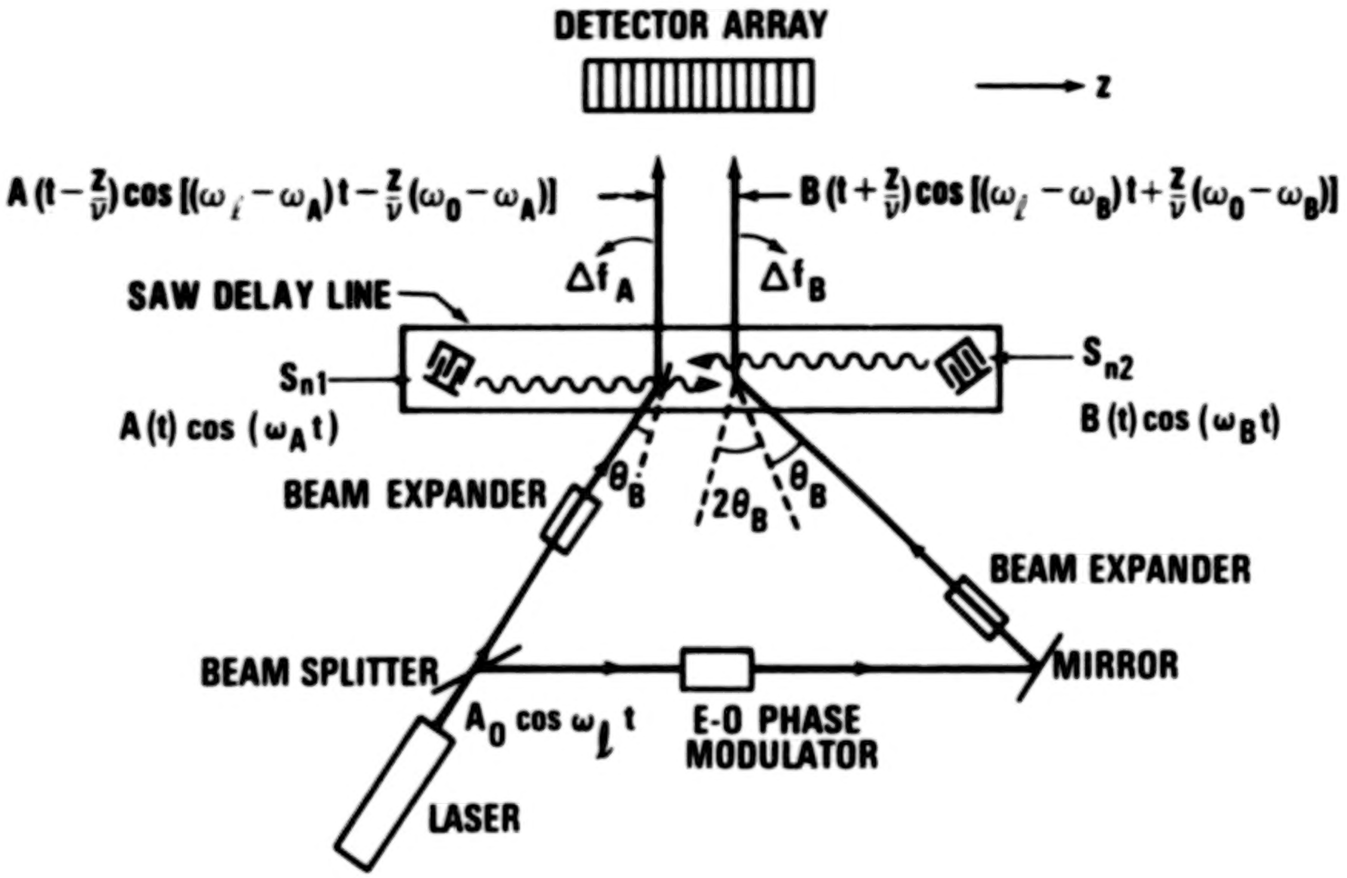

DETECTOR ARRAY
z
$A(t-\frac{z}{v})\cos[(\omega_\ell-\omega_A)t-\frac{z}{v}(\omega_0-\omega_A)]$
$B(t+\frac{z}{v})\cos[(\omega_\ell-\omega_B)t+\frac{z}{v}(\omega_0-\omega_B)]$
Δf_A
Δf_B
SAW DELAY LINE
S_{n1}
$A(t)\cos(\omega_A t)$
S_{n2}
$B(t)\cos(\omega_B t)$
BEAM EXPANDER
θ_B
$2\theta_B$
θ_B
BEAM EXPANDER
BEAM SPLITTER
MIRROR
$A_0\cos\omega_\ell t$
E-O PHASE MODULATOR
LASER

Figure 1.-Two-beam SAW time-integrating correlator.

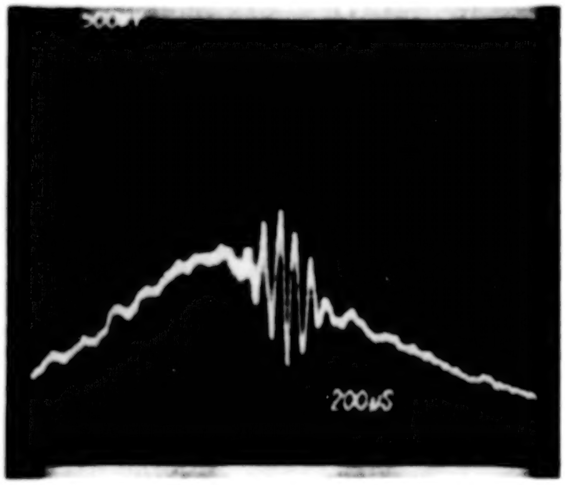


Figure 2.-Correlator output for direct sequence code at ω_0 = 1 Mhz.

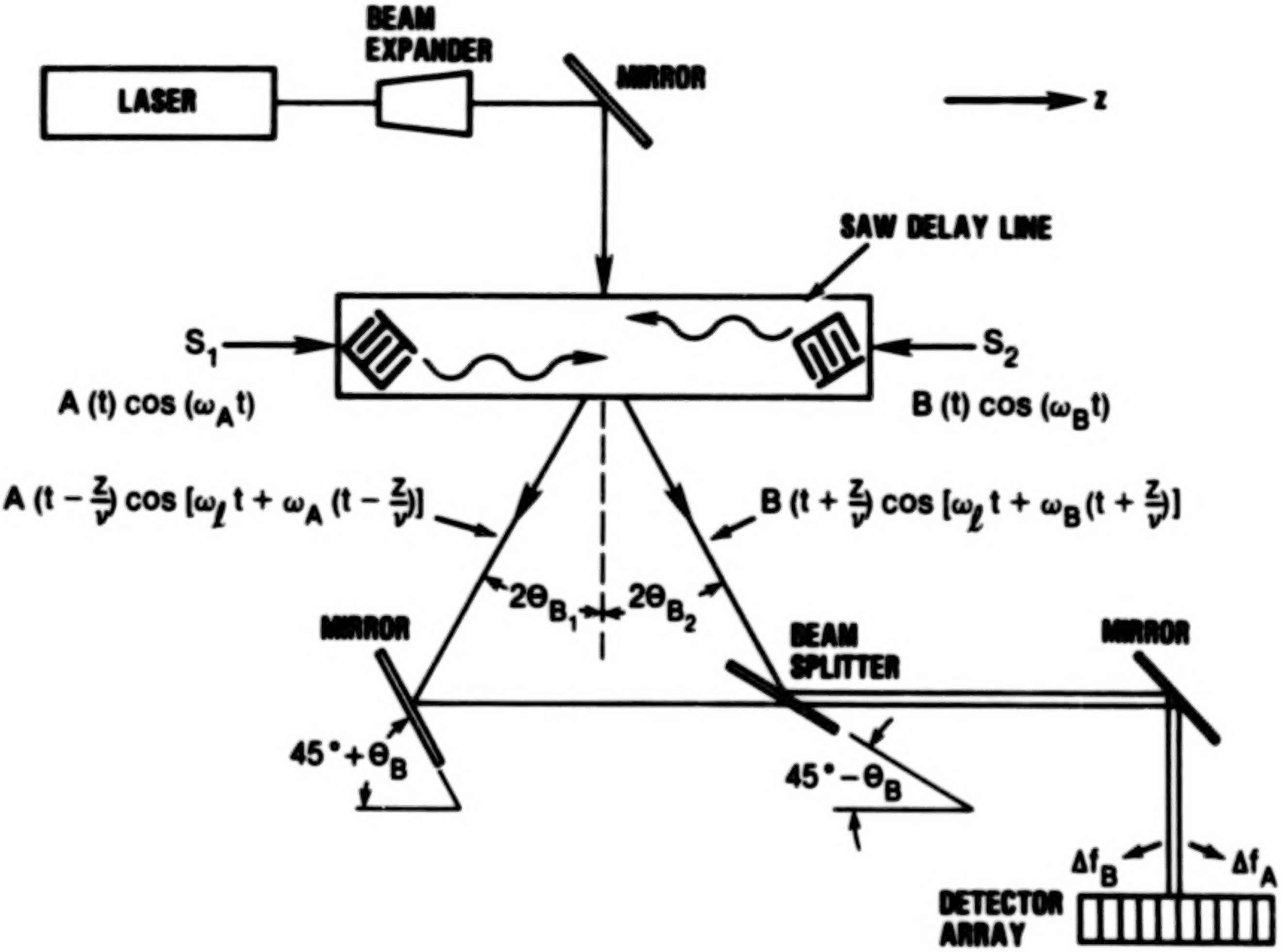
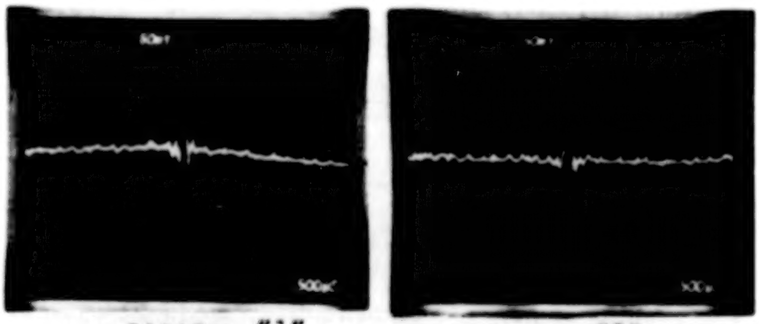

BEAM EXPANDER
LASER
MIRROR
$\rightarrow z$
SAW DELAY LINE
S_1
S_2
$A(t)\cos(\omega_A t)$
$B(t)\cos(\omega_B t)$
$A(t-\frac{z}{v})\cos[\omega_\ell t+\omega_A(t-\frac{z}{v})]$
$B(t+\frac{z}{v})\cos[\omega_\ell t+\omega_B(t+\frac{z}{v})]$
$2\Theta_{B_1}$
$2\Theta_{B_2}$
MIRROR
BEAM SPLITTER
MIRROR
$45°+\Theta_B$
$45°-\Theta_B$
Δf_B
Δf_A
DETECTOR ARRAY

Figure 3.-Single input beam SAW AO time-integrating correlator.



DIGITAL "1" DIGITAL "0"

Figure 4.- Output of correlator for hybrid bi-phase signals showing. demodulation.

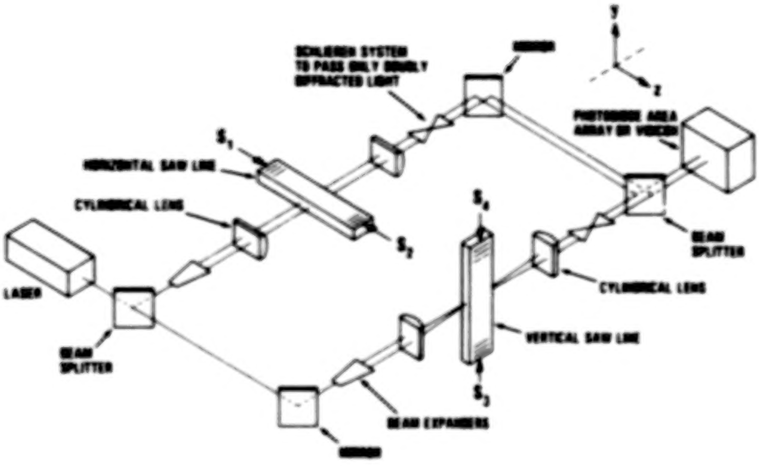

Y
Z
SCHLIEREN SYSTEM TO PASS ONLY DOUBLY DIFFRACTED LIGHT
MIRROR
PHOTODIODE AREA ARRAY OR VIDICON
S_1
HORIZONTAL SAW LINE
CYLINDRICAL LENS
S_2
S_4
BEAM SPLITTER
CYLINDRICAL LENS
LASER
BEAM SPLITTER
VERTICAL SAW LINE
S_3
BEAM EXPANDERS
MIRROR

Figure 5.-Two-dimensional SAW AO time-integrating correlator.

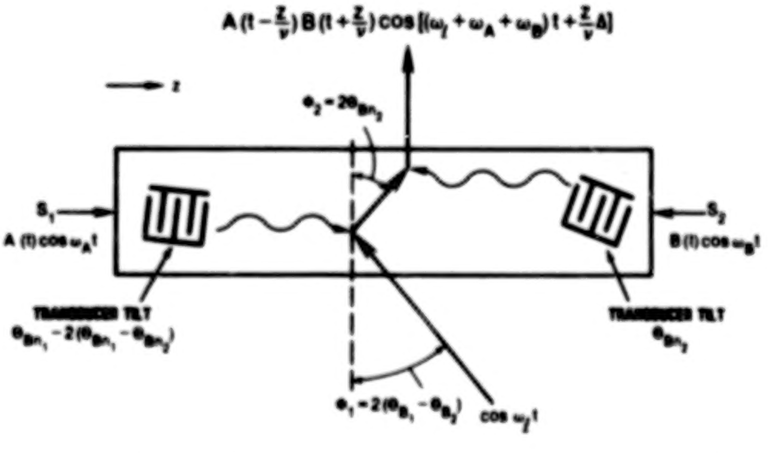

$A(t-\frac{Z}{v})B(t+\frac{Z}{v})\cos[(\omega_\ell+\omega_A+\omega_B)t+\frac{Z}{v}\Delta]$
$\rightarrow z$
$\theta_2 - 2\theta_{Bn_2}$
S_1
$A(t)\cos\omega_A t$
S_2
$B(t)\cos\omega_B t$
TRANSDUCER TILT
$\theta_{Bn_1} - 2(\theta_{Bn_1} - \theta_{Bn_2})$
TRANSDUCER TILT
θ_{Bn_2}
$\theta_1 - 2(\theta_{B_1} - \theta_{B_2})$
$\cos\omega_\ell t$

Figure 6.-Acousto-optic interactions producing a doubly diffracted beam.

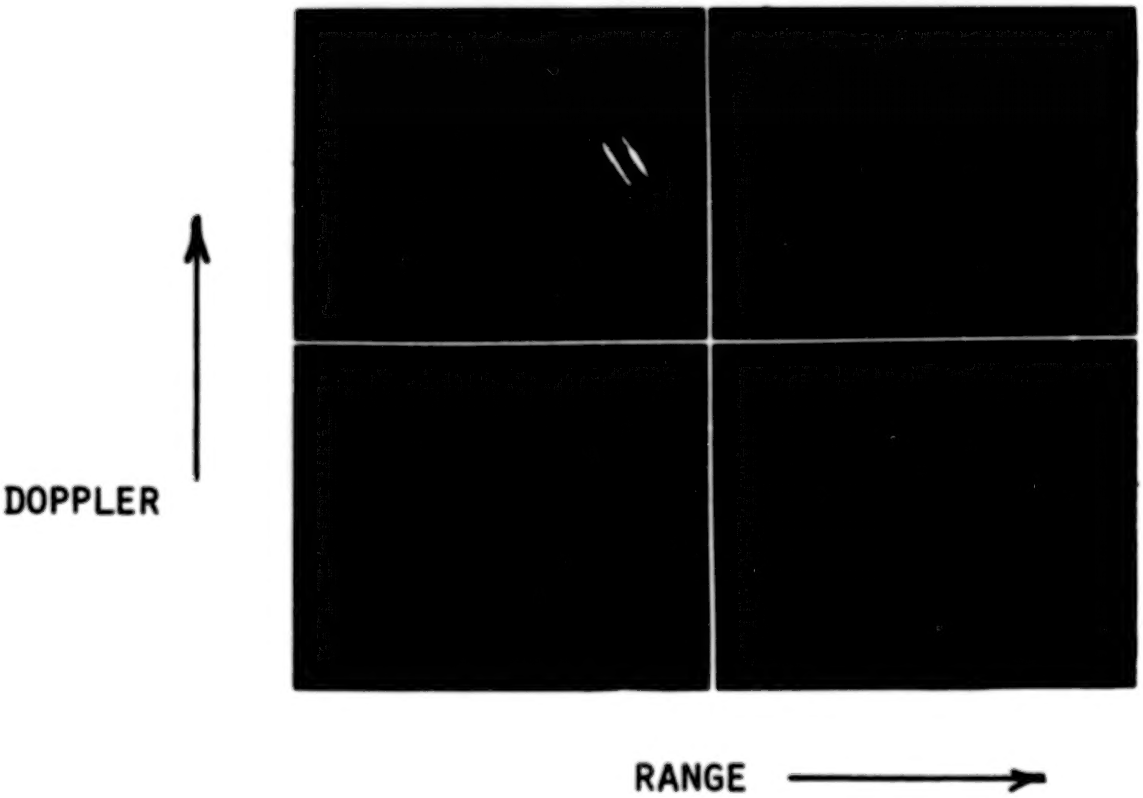

DOPPLER
RANGE

Figure 7.-Output of the four product correlator after bias subtraction.

NONCOHERENT OPTICAL PROCESSING OF RASTER FORMATTED SIGNALS*

William J. Miceli
Rome Air Development Center
Hanscom AFB, MA 01731

and

William W. Stoner
Science Applications, Inc.
Bedford, MA 01730

The inherent 2-D nature of optical imaging systems can be exploited in the processing of long duration signals by raster recording and parallel processing of complete frames of data. The 2-D Fourier transform of such a frame of data (i.e., the "folded spectrum") is most conveniently obtained with a coherent optical processor, however, the side effects of optical coherence (speckle, parasitic phase noise, linearity in complex amplitude, etc.) inhibit such an approach in many aerospace applications. These problems may be overcome by using noncoherent optical processors which use spatially incoherent light and capitalize upon the versatility of optical transfer function (OTF) synthesis. Incoherent optical spatial filtering of raster formatted signals is accomplished by imaging the raster recording with an optical system whose OTF modifies or selects out frequency components by means of a suitably designed pupil plane mask. To demonstrate the capability of incoherent optical signal processors, we present experimental results of a double cylindrical lens pupil mask which generates an OTF suitable for the selection of a single locus line in the folded spectrum. This OTF consists of three narrow passband ridges, with the center passband containing the dc component and the offset ones

---

* Because this paper was not available at the time of publication, only the abstract is included.

corresponding to the spatial frequency of the locus line of interest. The OTF is probed with a rotatable Ronchi ruling which is imaged onto a vidicon camera at the output of the system. Unless the spatial frequency and the angular orientation of the Ronchi ruling corrsponds to the frequency offset passbands of the OTF, the input is completely filtered out.

THREE METHODS FOR PERFORMING HANKEL TRANSFORMS

R. A. Athale, H. H. Szu and J. N. Lee
Naval Research Laboratory, Washington, D. C. 20375

ABSTRACT

Generalized Hankel transforms are useful in analyzing the effect of circularly symmetric optical systems on arbitrary inputs. Some examples of such systems are complex laser resonators and space telescopes. Three methods for performing Hankel transforms with optical or digital processors are described. The first method is applicable when the input data is available in Cartesian (x-y) format and uses the close connection between generalized Hankel transform and the two-dimensional Fourier transform in Cartesian coordinates. The second method is useful when the input data is in polar (r - θ) format and uses change of variables to perform the n^{th} order Hankel transform as a correlation integral. The third method utilizes the von Neumann addition theorem for Bessel functions to extract the Hankel coefficients from a correlation between the radial part of the input and a Bessel function. Initial experimental results obtained for optical implementation of the first two methods are presented.

INTRODUCTION

The analysis of complex optical systems is greatly facilitated by two-dimensional Fourier transform techniques. The effect of an optical system on arbitrary inputs is easily described by a transfer function in the Fourier domain. Generalized Hankel transforms are similarly useful when dealing with a circularly symmetric (or axisymmetric) system for arbitrary inputs.[1] This situation is encountered in performing mode analysis on the output of a slightly misaligned laser resonator as well as in aligning space telescopes. An optical method for performing mode analysis via generalized Hankel transform will have the unique advantage of preserving the phase of the wavefront to be analyzed.

It is well known that when a two-dimensional function has circular symmetry (i.e. it depends only on the radial variable, r), its Fourier transform is also circularly symmetric (i.e. it depends only on the radial variable, ρ). It can be shown that in such a case the Fourier transform is equivalent to the 0^{th} order Hankel transform.[2]

If

$$f(x,y) \equiv f\left(\sqrt{x^2+y^2}\right)$$

Then

$$\mathscr{F}_2\left\{ f\left(\sqrt{x^2+y^2}\right)\right\} = \mathscr{F}\left(\sqrt{u^2+v^2}\right)$$

where $\mathscr{F}_2\{\ \}$ indicates two-dimensional Fourier transformation.

For

$$r = \sqrt{x^2+y^2}\ , \ \rho = \sqrt{u^2+v^2}$$

$$\mathscr{F}\left(\sqrt{u^2+v^2}\right) \equiv F_o(\rho) = 2\pi \int_o^\infty rdr\ f(r)\ J_o(\rho) \qquad (1)$$

where $F_o(\rho)$ is the 0^{th} order Hankel transform and $J_o(\rho r)$ is the 0^{th} order Bessel function of the first kind. Thus in dealing with circularly symmetric systems, the 0^{th} order Hankel transform (which is a one-dimensional operation) can be used instead of the two-dimensional Fourier transform if the inputs also are circularly symmetric. For arbitrary inputs, this technique can be extended by using the generalized Hankel transform, which expands on the n^{th} order transform ($F_n(\rho)$ with kernel $J_n(\rho)$ in Eq. (1)).

The generalized Hankel transform, $F_{nn}(\rho)$, can be defined for an arbitrary function, $f(r,\theta)$, as follows:

$$f(r,\theta) = \sum_{n=-\infty}^{\infty} f_n(r)\ e^{jn\theta}$$

$$F_{nn}(\rho) = \int_o^\infty rdr\ f_n(r)\ J_n(\rho r) \qquad (2)$$

The generalized Hankel transform thus involves a Fourier series expansion in θ, followed by an n^{th} order Hankel transform (with n^{th} order Bessel function, $J_n(\rho r)$, as the transformation kernel) on the n^{th} coefficient of expansion, $f_n(r)$. This generalized Hankel transform is useful in analyzing systems with circular symmetry but when the input is not circularly symmetric.

In the following sections we will describe three techniques for easy implementation of the generalized Hankel transform with optical or digital processors. These three techniques are each applicable in different circumstances. Initial experimental results on the optical implementation of the first two methods will be presented. In conclusion, we will detail the course of future work in this area.

OPTICAL IMPLEMENTATION OF GENERALIZED HANKEL TRANSFORM

In optical processors a two-dimensional Fourier transform with respect to the Cartesian coordinates (x,y) is performed very easily with the help of a simple spherical lens.[2] The equivalence of two-dimensional Fourier transform and 0^{th} order Hankel transform for circularly symmetric functions was established in the Introduction. A similar connection exists between two-dimensional Fourier transform with respect to Cartesian coordinates (x,y) and n^{th} order Hankel transform of the radial part if the function is of a form $f(r)\ e^{jn\theta}$.

$$\mathscr{F}_2\left\{ f(r)\, e^{jn\theta} \right\} = 2\pi F_n(\rho)\, e^{-jn\phi} \tag{3}$$

where

$$\rho = (u^2+v^2)^{\frac{1}{2}}\ ,\ \phi = \tan^{-1}(v/u).$$

and by definition

$$F_n(\rho) = \int_0^\infty r dr\, f(r)\, J_n(\rho r)\ .$$

This connection arises out of the integral representation of $J_n(\rho r)$.[1] Using this result in the definition of generalized Hankel transform given in Eq. (2) the following relation is obtained between the two-dimensional Fourier transform and the generalized Hankel transform:

$$\mathscr{F}_2\left\{ f(r,\theta) \right\} = \mathscr{F}_2\left\{ \sum_{n=-\infty}^{\infty} f_n(r)\, e^{jn\theta} \right\}$$

$$= 2\pi \sum_{n=-\infty}^{\infty} F_{nn}(\rho)\, e^{-jn\phi} \tag{4}$$

Thus the generalized Hankel transform $F_{nn}(\rho)$ is equivalent to the n^{th} coefficient of Fourier series expansion of $\mathscr{F}(\rho,\phi)$ in variable ϕ. Using this result, an optical system shown in Fig. 1 is designed to perform generalized Hankel transform on an arbitrary input that is available in Cartesian format $(f(x,y))$. The spherical lens performs a two-dimensional Fourier transform on $f(x,y)$ generating $\mathscr{F}(u,v)$. A suitably designed computer generated hologram then performs the Cartesian-to-polar $(u,v \rightarrow \rho,\phi)$ coordinate transformation on $\mathscr{F}(u,v)$ generating $\mathscr{F}(\rho,\phi)$.[3] This is followed by a cylindrical lens which takes a one-dimensional Fourier transform with respect to variable ϕ generating the desired output, $F_{nn}(\rho)$.

Initial optical experiments established the connection between two-dimensional Fourier transform and n^{th} order Hankel transform for a function of the form $f(r)e^{jn\theta}$. The schematic diagram of the optical system is shown in Fig. 2. A computer generated hologram was used to encode $e^{jn\theta}$ dependence of the input. The output was detected by a TV camera, which measures the light intensity in the Fourier plane of the input. The input used in these experiments had an r dependence given by $\delta(r-a)$, thus corresponding to a thin ring of radius "a" in the Cartesian (x,y) plane. The n^{th} order Hankel transform of $\delta(r-a)$ is $aJ_n(a\rho)$, making the output easily understandable. Figure 3 shows the results corresponding to 0^{th} order Hankel transform (i.e. $e^{jn\theta} = 1$). In Fig. 3(b) the function $|J_o(a\rho)|^2$ is plotted, which is then compared to a line scan through the origin (Fig. 3(c)) of the output shown in Fig. 3(a). Very good qualitative agreement between the theoretical and experimental results is obtained. Figures 4(a-c) present the results for 1^{st} order Hankel transform for the same f(r), but here the θ dependence is $e^{j\theta}$. Again good qualitative agreement is seen between theory and experiment. A line scan through the origin of the output corresponding to 2^{nd} order Hankel transform of $\delta(r-a)$ (with θ dependence $e^{j2\theta}$) is shown in Fig. 5. The zero at the origin was broader and the side lobes were seen to fall slower indicating that we indeed have $|J_2(a\rho)|^2$ as expected.

OPTICAL IMPLEMENTATION OF N^{TH} ORDER HANKEL TRANSFORM

The previous method is applicable when the input is available in Cartesian (x,y) format, since it involves performing two-dimensional Fourier transforms with respect to the Cartesian variables. If the input is polar (r,θ) formatted, a more direct approach outlined in Eq. (3) has to be followed in obtaining generalized Hankel transform. The first part of the operation, which involves a Fourier series expansion in variable θ, is easily performed optically using a cylindrical lens. The calculation of n^{th} order Hankel transform of the n^{th} coefficient of expansion is less straightforward since it corresponds to a space-variant operation. So the main aim of the next two methods is to perform the space-variant operation of n^{th} order Hankel transform optically.

a. Method Using Change of Variables

One standard procedure used in converting a space-variant operation into a shift-invariant operation is to employ appropriate change of variables. In the case of n^{th} order Hankel transforms the following procedure was described by Siegman for implementing the space-variant operation as a correlation integral on a digital processor.[4] From the definition

$$F_n(\rho) = \int_0^\infty r dr\, f(r)\, J_n(\rho r) ,$$

using

$$r = r_o\, e^{\alpha x},\ \rho = \rho_o e^{\alpha y} \tag{5}$$

$$\hat{F}_n(y) = \int_{-\infty}^{\infty} \hat{f}(x)\, \hat{J}_n(x+y)\, dx$$

where $\hat{F}_n(y) = \rho F_n(\rho)$, $\hat{f}(x) = rf(r)$, and $\hat{J}_n(x+y) = \alpha r\rho\, J_n(r\rho)$. The algorithm, therefore, consists of first linearly weighting the input f(r) and performing $r\rightarrow x$ (logarithmic related) coordinate transform. This distorted input is then correlated with a similarly weighted and coordinate-transformed n^{th} order Bessel function to give the desired Hankel transform as well as linearly weighted and coordinate distorted forms. In any physical system the correlation integral will be performed over a finite interval, giving rise to truncation errors. Also if the input is sampled in the x-domain, the sampling rate should be adequate to represent the function accurately in x-domain.

These factors and others are discussed at length, especially for digital implementation, in Refs. 4 and 5.

Since the operations of coordinate transformation and correlation can be performed by an optical processor, the optical system outlined in Fig. 6 can calculate nth order Hankel transforms. Computer generated holograms are used to perform r→x coordinate transformation as well as to encode the Fourier plane filter with impulse response $\hat{J}_n(x)$. The second dimension of the optical processor can be used to perform different order Hankel transforms on different inputs, thus achieving multichannel operation.

In the initial experiment, both the input and the Fourier plane filter were encoded by computer generated holograms. The linear weighting and the coordinate transformation of the input was performed by the digital computer before doing the holographic encoding. The computer generated holograms used the Lee-Burckhardt technique[6] and contained 128 pixels. The optical system is depicted in Fig. 7. This system was then used to perform 0th order Hankel transform on two different inputs, $f_1(r) = \rho_1 J_0(\rho_1 r)$ and $f_2(r) = \rho_2 J_0(\rho_2 r)$. The result of 0th order Hankel transform on $f_1(r)$ and $f_2(r)$ should be $\delta(\rho-\rho_1)$ and $\delta(\rho-\rho_2)$ respectively. The results of the computer simulation of this algorithm are depicted in Fig. 8. The finite width of the peak and the sidelobes are due to the finite limits of intergration. To facilitate easy comparison with the experimental results, the 0th order Hankel transform with linear weighting and coordinate distortion (i.e. $|\hat{F}_0(y)|^2$) was plotted versus y instead of $F_0(\rho)$ versus ρ. Figure 9 shows the experimental results obtained. The optical output was detected by a 1024-element Reticon linear photodiode array. The shift in the peak position and the difference in the peak heights (due to the linear weight) are evident, indicating good qualitative agreement with the computer simulation results.

b. Method Using Neumann Addition Theorem

This method investigates an approach based on the special properties of Bessel functions.[7] If the input f(r) is correlated with $J_m(\rho r)$ then the mth order Hankel coefficient $F_m(\rho)$ is obtained at the origin of the correlation plane for the particular value of ρ encoded in the Bessel function kernel.

$$F_m(\rho) = \int_0^\infty r\,dr\; f(r)\; J_m(\rho(r+r'))\Big|_{r'=0} \qquad (6)$$

The Neumann addition theorem for Bessel functions states that

$$J_m(\rho(r+r')) = \sum_{n=-\infty}^{\infty} J_{m-n}(\rho r')\; J_n(\rho r) \qquad (7)$$

Substituting for $J_m(\rho(r+r'))$ from Eq. (7) into Eq. (6) we get

$$\int_0^\infty r dr\ f(r)\ J_m(\rho(r+r')) = \int_0^\infty r dr\ f(r) \sum_{n=-\infty}^{\infty} J_{m-n}(\rho r')\ J_n(\rho r) \tag{8}$$

$$= \sum_{n=-\infty}^{\infty} J_{m-n}(\rho r')\ F_n(\rho).$$

Thus it is seen that the correlation plane contains an infinite sum of Bessel functions of different order weighted by Hankel transform coefficients of different order. So in principle it is possible to extract Hankel transform coefficients of different order out of a single one-dimensional correlation operation.

The other dimension of optical system can be used to perform correlations with $J_m(\rho r)$ with different values of ρ to obtain complete Hankel transform. This approach is currently being evaluated further to determine the situations in which this method will be suitable.

FUTURE WORK

In this paper we have presented brief outlines of three different approaches for performing generalized Hankel transforms by optical or digital processors. We also presented initial experimental results on two of the three approaches. The future work will concentrate on carrying out the coordinate transformation required by the first two methods optically via computer generated holograms. A more quantitative analysis of the performance of the optical processor will be carried out. These three methods will be compared with each other with respect to space-bandwidth requirements as well as pre- and post-processing requirements. Finally, generalized Hankel transforms will be applied to specific problems, such as analysis of complex laser resonators, and optical systems will be developed for those problems.

ACKNOWLEDGEMENT

We wish to thank Dr. J. R. Leger for producing the computer generated holograms using the facilities at University of California, San Diego.

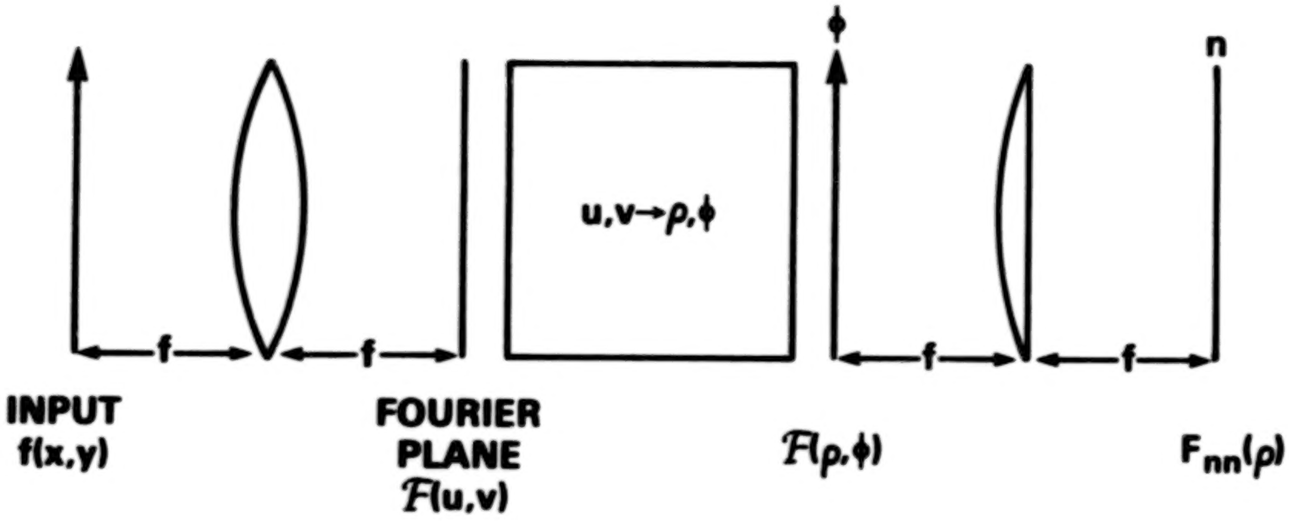
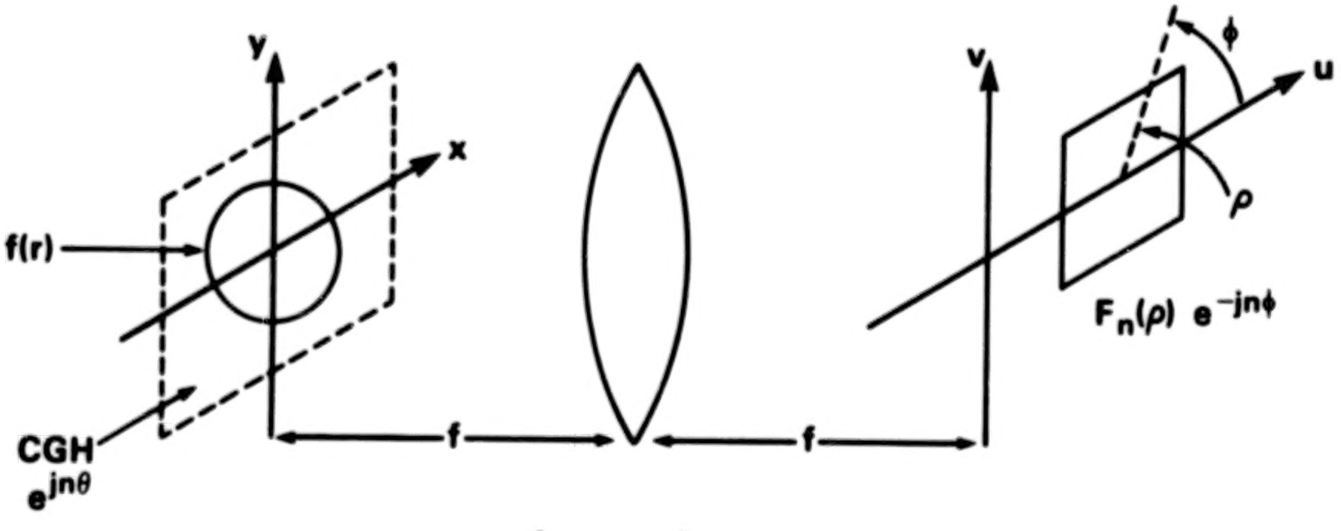
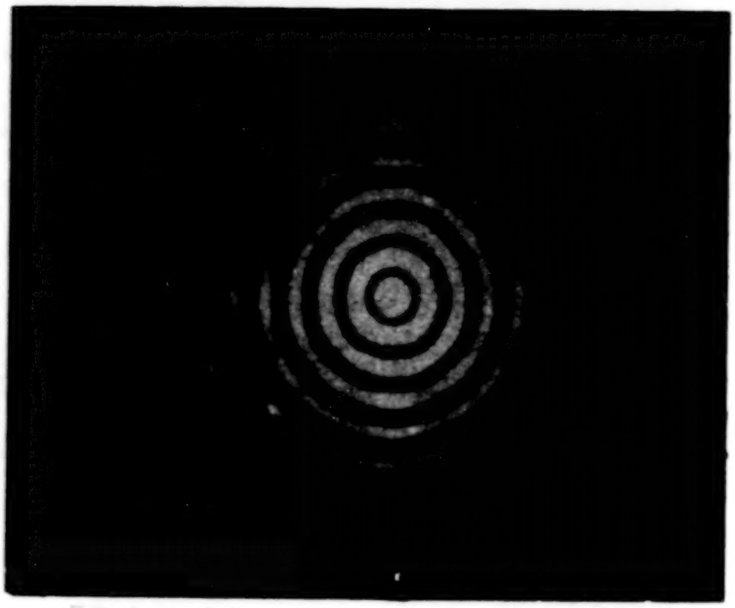

ϕ
n
u,v→ρ,ϕ
f
f
f
f
INPUT
f(x,y)
FOURIER
PLANE
$\mathcal{F}(u,v)$
$\mathcal{F}(\rho,\phi)$
$F_{nn}(\rho)$

Figure 1. - Schematic diagram of optical processor for performing generalized Hankel transform via a two-dimensional Fourier transform with respect to Cartesian coordinates (x,y).

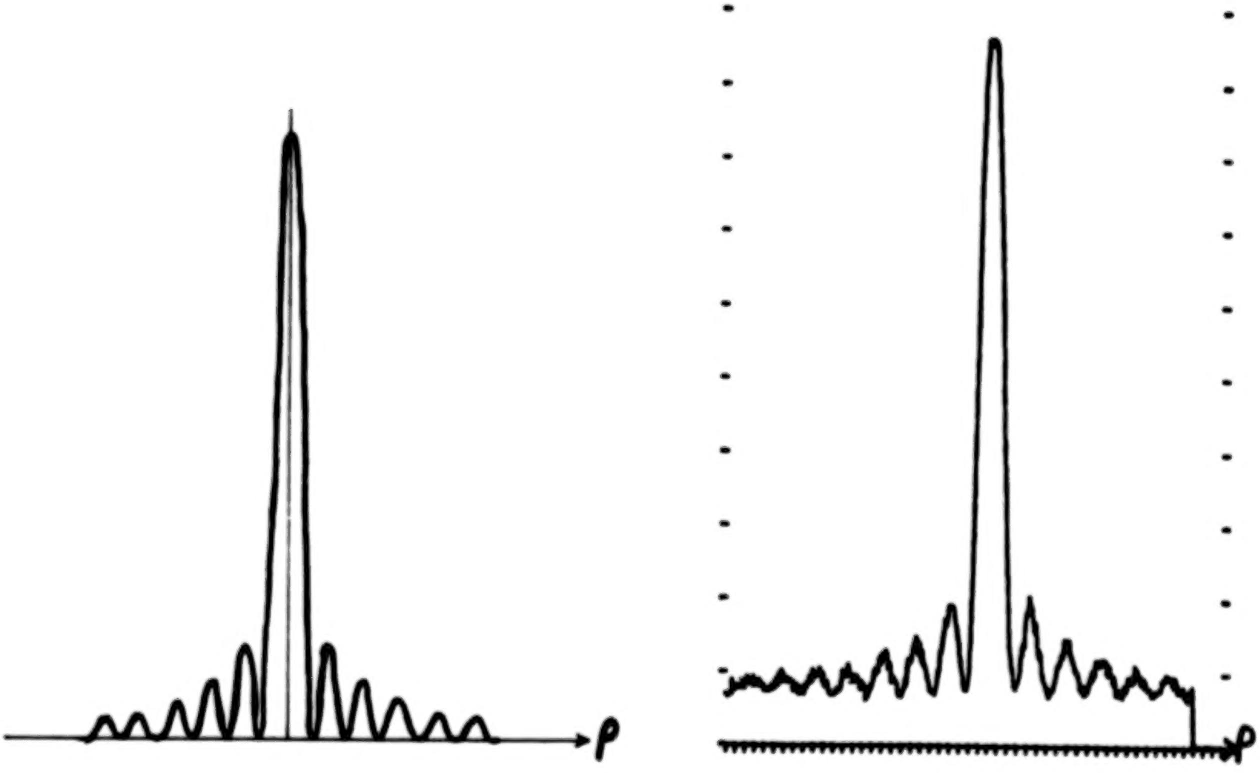

y
x
f(r)
CGH
$e^{jn\theta}$
f
f
v
u
ϕ
ρ
$F_n(\rho)\ e^{-jn\phi}$

$$\mathcal{F}_2 \left\{ f(r)\ e^{jn\theta} \right\} = F_n(\rho)\ e^{-jn\phi}$$

Figure 2. - Schematic diagram of experimental setup for establishing the relation between n^{th} order Hankel transform of f(r) and two-dimensional Fourier transform of f(r) $e^{jn\theta}$ with respect to Cartesian coordinates (x,y).



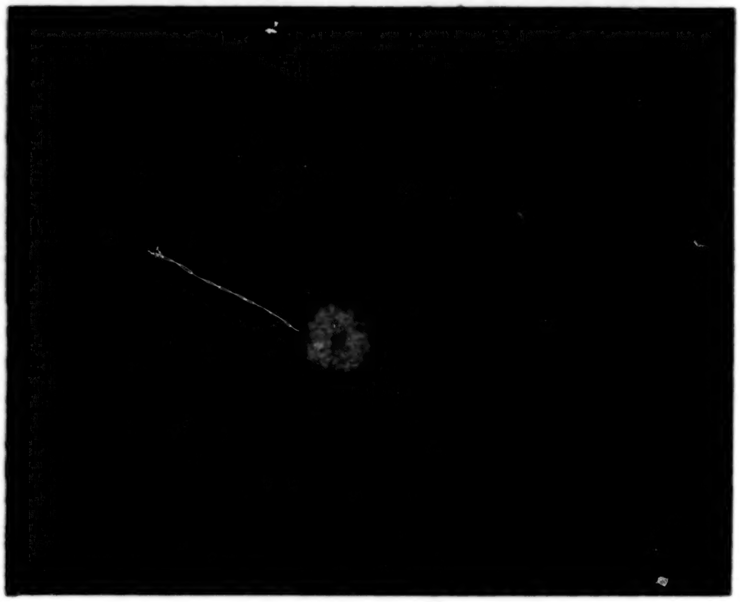
(a) Photograph of the output.



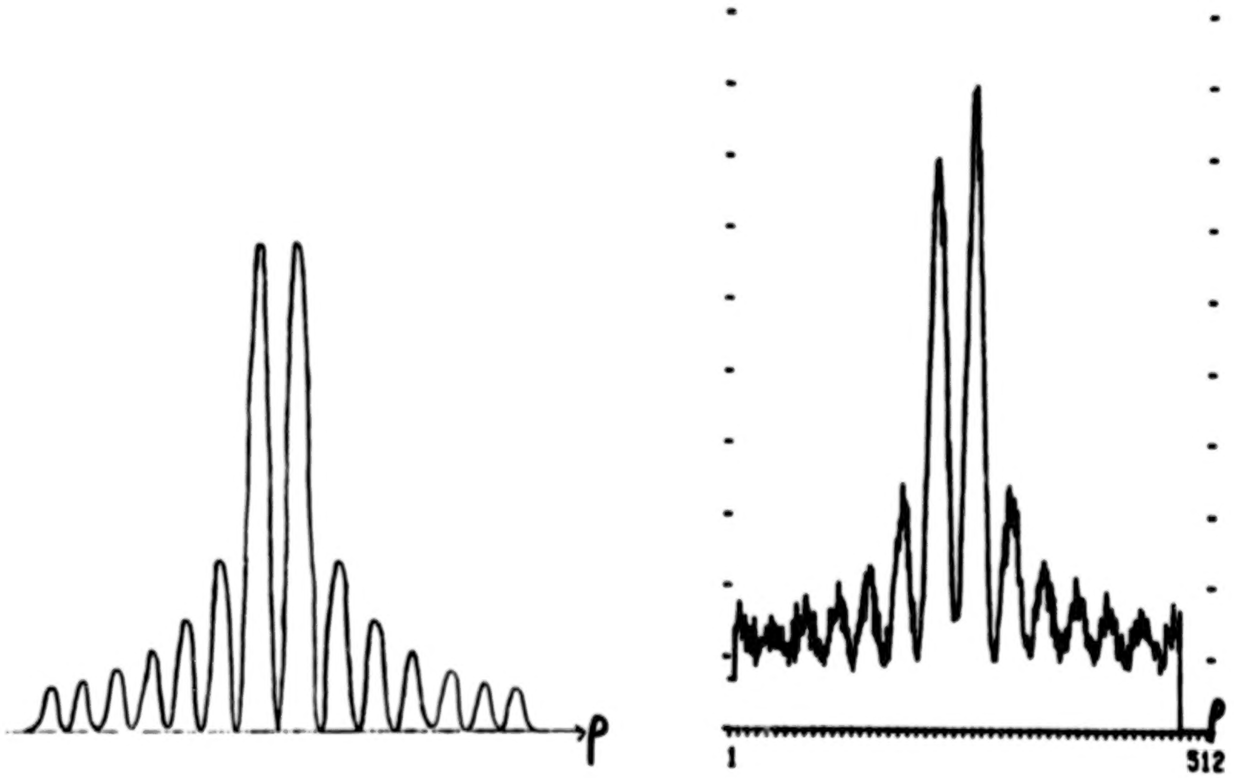
(b) Plot of $|J_o(a\rho)|^2$ versus ρ, which is the theoretically expected result.

(c) Line scan through the origin of the pattern in figure 3 (a), giving $|F_o(\rho)|^2$ versus ρ.

Figure 3. - Experimental results for 0^{th} order Hankel transform $F_o(\rho)$ of $\delta(r-a)$.



(a) Photograph of the output.



(b) Plot of $|J_o(a\rho)|^2$ versus ρ, which is the theoretically expected result.

(c) Line scan through origin of the bottom in figure 4 (a), giving $|F_1(\rho)|^2$ versus ρ.

Figure 4. - Experimental results for 1st order Hankel transform $F_1(\rho)$ of $\delta(r-a)$.

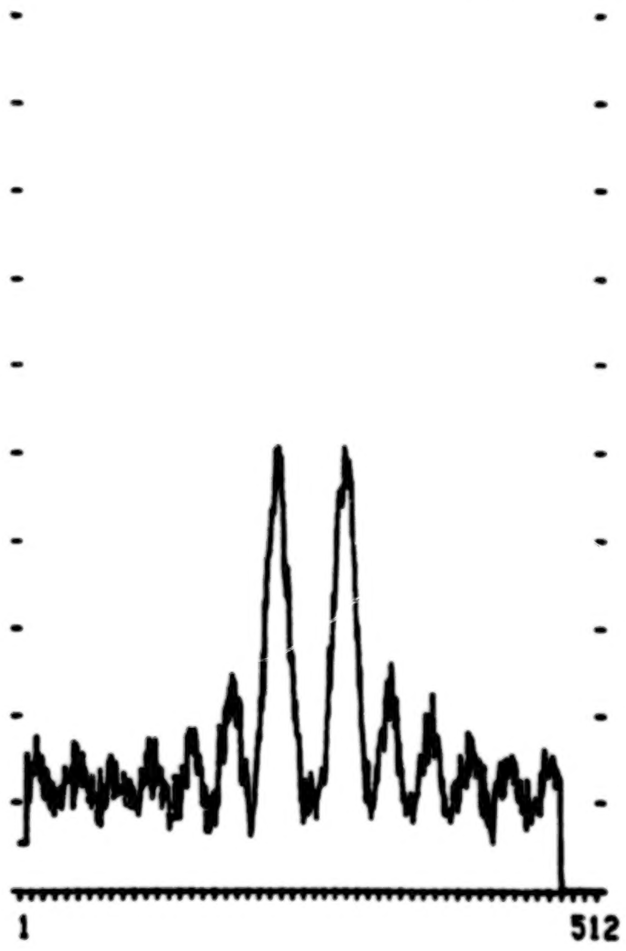


Figure 5. - Line scan through the origin of the two-dimensional Fourier transform of $\delta(r - a)\, e^{j2\theta}$, giving $|F_2(\rho)|^2$ for $\delta(r - a)$.

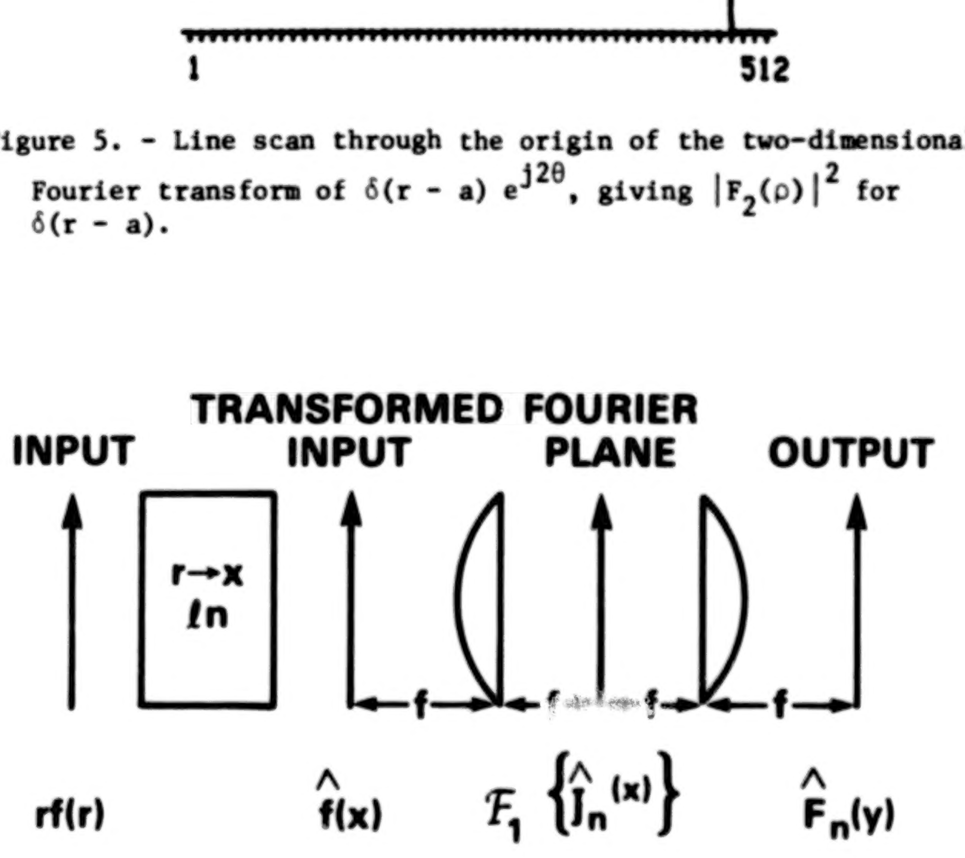

TRANSFORMED INPUT
FOURIER PLANE
INPUT
OUTPUT
r→x
ln
f
f
f
f
rf(r)
$\hat{f}(x)$
$\mathcal{F}_1 \{\hat{J}_n(x)\}$
$\hat{F}_n(y)$

Figure 6. - Schematic diagram of an optical processor for performing n^{th} order Hankel transform on f(r) employing change of variables.

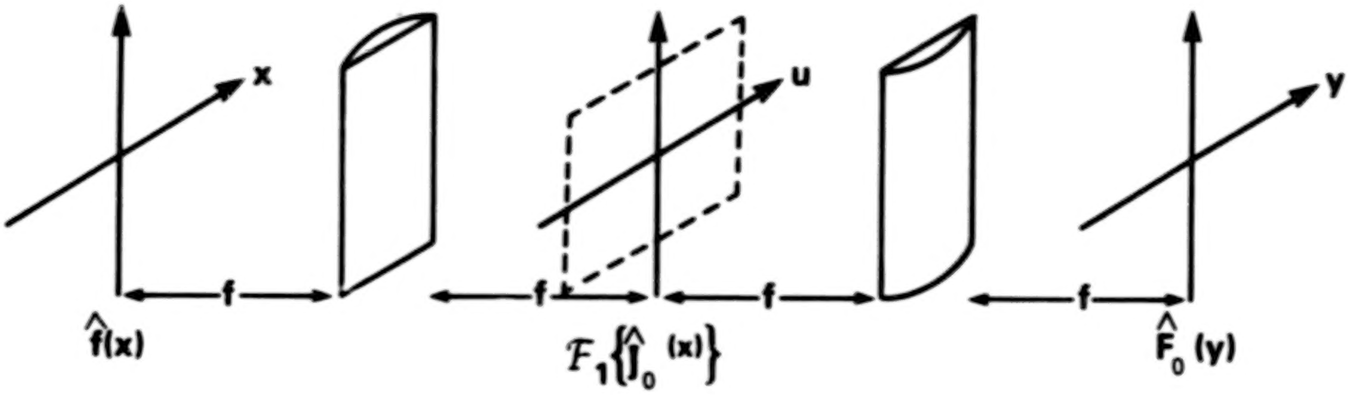


Figure 7. - Schematic diagram of the experimental setup for performing the one-dimensional correlation between linearly weighted and coordinate transformed input, $\hat{f}(X)$, and similarly weighted and transformed Bessel function $\hat{J}_o(X)$.

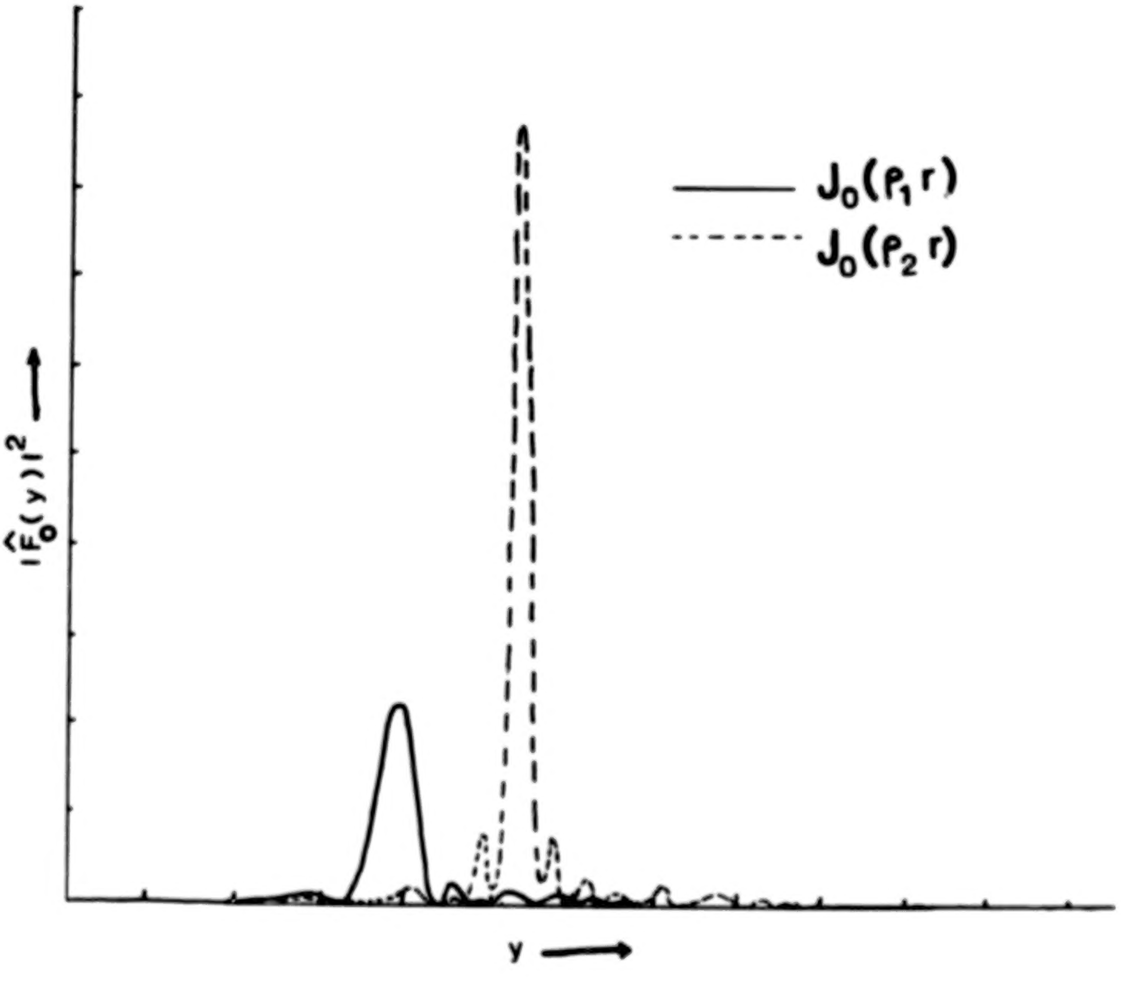
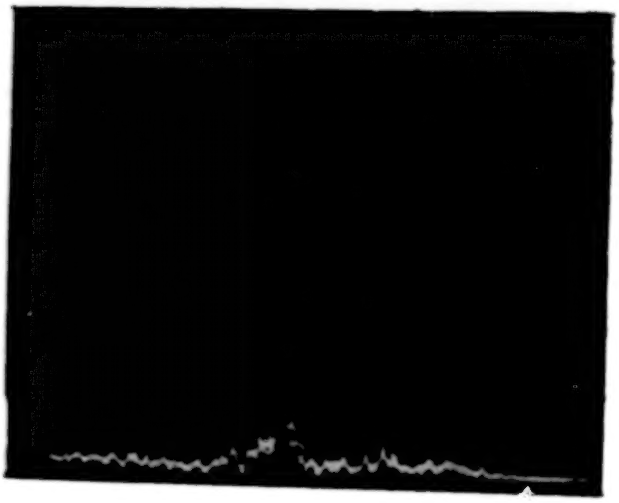
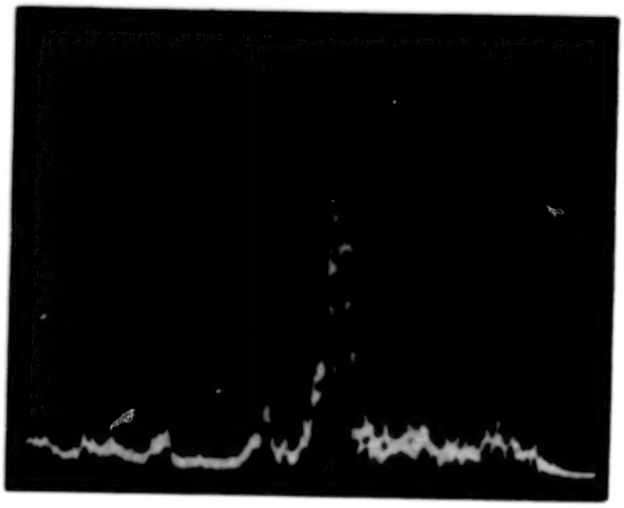


Figure 8. - Results of computer simulation of 0^{th} order Hankel transform of $\rho_1 J_o(\rho,r)$ and $\rho_2\ J_o(\rho_2 r)$. The linearly weighted and coordinate transformed Hankel coefficients, $|\hat{F}_o(y)|^2$, are plotted versus y. $\rho_2 = 2\rho_1$.



(a) Input $f_1(r) = \rho_1 \; J_o(\rho_1 r)$.



(b) Input $f_2(r) = \rho_2 \; J_o(\rho_2 r)$.

Figure 9. - Oscilloscope traces of output of the optical processor performing the 0^{th} order Hankel transform. Traces correspond to $|\hat{F}_o(y)|^2$ versus y for the two inputs.

OPTICAL RECOGNITION OF STATISTICAL PATTERNS

Sing H. Lee
Electrical Engineering & Computer Sciences
University of California, San Diego 92093

SUMMARY

This paper describes the optical implementation of the Fukunaga-Koontz transform (FKT) and the Least-Squares Linear Mapping Technique (LSLMT). The FKT is a linear transformation which performs image feature extraction for a two-class image classification problem. It has the property that the most important basis functions for representing one class of image data (in a least-squares sense) are also the least important for representing a second image class. The LSLMT is useful for performing a transform from large dimensional feature space to small dimensional decision space for separating multiple image classes by maximizing the interclass differences while minimizing the intraclass variations. The FKT and the LSLMT were optically implemented by utilizing a coded phase optical processor. Good experimental results were obtained, and they were compared with the performance of the matched filter and the average filter.

I. INTRODUCTION

Optical matched filtering, which is invariant to the translation of input, has been the main basis for optical pattern recognition for many years (refs. 1-3). Optical Mellin transform was later introduced to obtain scale invariant correlation (ref. 4). Currently, the subject of performing statistical pattern recognition and classification optically is found to be of considerable interest (refs. 5-9). For the two-class problem, we shall show in this paper that the Fukunaga-Koontz transform (FKT) can be implemented optically (ref. 9). For the K-class problem where K is greater than two, the optical implementation of the Least-Squares Linear Mapping Technique (LSLMT) will be presented (ref. 10).

Figure 1 shows the hybrid system for optical implementation of FKT and LSLMT. The key element in the hybrid system is the computer generated spatial filter. The filter is synthesized from K training sets of images using the FKT or LSLMT algorithm. When the LSLMT filter is inserted in the filter plane and a test image in the input plane of a coherent optical processor (COP), the output from the COP will be a bright spot of light in one of K predetermined locations, provided that the test image belongs to one of the K image classes in the statistical sense and it is illuminated with a random phase wavefront. As fast as one can input a new test image through a real-time interface device to the COP, a new bright spot of light indicating classification will appear among the K predetermined locations in the output. The COP effectively performs a matrix-vector multiplication in real time, where the matrix is $(K \times N)$, the vector is $(N \times 1)$, $N(= n \times n)$ is the space-bandwidth product or the number of pixels in an image, and K is the number of image classes.

When the FKT filter (instead of the LSLMT filter) is used, the output from the COP will contain a number of light spots whose intensities correspond to the squares of FKT coefficients. The FKT coefficients are associated with basis functions (features) which possess the interesting property that the most important basis function for one class is also the least important basis function of the other class. Image classification can become as simple as comparing two FKT coefficients associated with the two basis functions of high separating power.

II. THE FUKUNAGA-KOONTZ TRANSFORM

A. Procedures for Calculating the Basis Function

The procedures for calculating the basis functions of the F-K transform are summarized as follows. (Details can be found in ref. 9.)

(a) Given a training set of images $\{\underline{X}_j^{(i)}\}$ consisting of M_1 image samples from class 1 ($i = 1$; $j = 1, \ldots, M_1$) and M_2 samples from class 2 ($i = 2$; $j = 1, \ldots, M_2$), we first represent each sample image as a column vector of length N, and define matrices W_1, W_2 and W_t as follows:

$$W_1 = \left[\underline{X}_1^{(1)}\,\underline{X}_2^{(1)}\,\underline{X}_3^{(1)} \cdots \underline{X}_{M_1}^{(1)}\right], \tag{1}$$

$$W_2 = \left[\underline{X}_1^{(2)}\,\underline{X}_2^{(2)}\,\underline{X}_3^{(2)} \cdots \underline{X}_{M_2}^{(2)}\right],$$

and

$$W_t = \left[\underline{X}_1^{(1)}\,\underline{X}_2^{(1)} \cdots \underline{X}_{M_1}^{(1)} \vdots\, \underline{X}_1^{(2)}\,\underline{X}_2^{(2)} \cdots \underline{X}_{M_2}^{(2)}\right]$$

where W_1 is an $(N \times M_1)$ matrix, W_2 is $(N \times M_2)$ and W_t is $(N \times (M_1 + M_2))$ containing sample images from both classes. The sample correlation matrix of the whole process (computed with data from both classes) is then given by

$$W_t W_t^+ = W_1 W_1^+ + W_2 W_2^+ , \tag{2}$$

since

$$W_i W_i^+ = \sum_{j=1}^{M_i} \underline{X}_j^{(i)} \underline{X}_j^{(i)^+} , \quad i = 1, 2$$

is the sample correlation matrix for the i^{th} class. Finding the basis functions normally involves diagonalizing the $(N \times N)$ matrix of $(W_t W_t^+)$, which is a very time consuming task. But, when the intra-class variation in each class is small, this time consuming task can be simplified to the following steps.

(b) We find the $(M_1 + M_2)$ eigenvectors E and eigenvalues Λ of $(W_t^+ W_t)$,

$$(W_t^+ W_t)E = E\Lambda. \tag{3}$$

It can be seen from multiplying Eq. (3) from the left on both sides by W_t and regrouping that diagonalizing $(W_t W_t^+)$ gives $(W_t E)$ and Λ as the eigenvectors and eigenvalues

$$(W_t W_t^+)(W_t E) = (W_t E)\Lambda \tag{4}$$

Therefore, the time consuming task of diagonalizing $(W_t W_t^+)$ has become the much easier problem of diagonalizing the $[(M_1 + M_2) \times (M_1 + M_2)]$ matrix $(W_t^+ W_t)$ and calculating the product $(W_t E)$.

(c) By choosing the M_1 eigenvectors with the largest eigenvalues, we construct matrices E_c and Λ_c of dimensions $[(M_1 + M_2) \times M_1]$ and $[M_1 \times M_1]$, respectively.

(d) The eigenvectors Ψ, Θ and eigenvalues Γ, M of the transformed single class correlation matrices are given by

$$\left[\Lambda_c^{-1} E_c^{+} W_t^{+} (W_1 W_1^{+}) W_t E_c \Lambda_c^{-1}\right]\Psi = \Psi\Gamma \tag{5a}$$

$$\left[\Lambda_c^{-1} E_c^{+} W_t^{+} (W_2 W_2^{+}) W_t E_c \Lambda_c^{-1}\right]\Theta = \Theta M \tag{5b}$$

It is shown in reference 9 that the eigenvectors of equations (5a) and (5b) are identical,

$$\Theta = \Psi \tag{6a}$$

and that the eigenvalues of equations (5a) and (5b) are related by

$$M = 1 - \Gamma \tag{6b}$$

Thus, the eigenvectors with large eigenvalues applied to one class will have small eigenvalues when applied to the other class.

(e) The basis functions for the F-K transform are $(\psi_i^{+} \Lambda_c^{-1} E_c^{+} W_t^{+})$, where ψ_i are the eigenvectors in Ψ.

B. Optical Implementation of the F-K Transform

Since the F-K transform is a linear transformation, the coefficient corresponding to a specific basis function is found by calculating the inner product between an input image function and the complex conjugate of the basis function. A coherent optical processor can be designed to calculate these coefficients in parallel by multiplying the input image by a coded phase function $\exp[i\phi_r(x,y)]$, and designing a filter whose impulse response $h^*(-x,-y)$ consists of the complex conjugate of a summation of shifted products of the coded phase function and a particular basis function

$$h^*(-x,-y) = \sum_{pq}^{M} f_{pq}^*(x + p\Delta,\ y + q\Delta)\exp[-i\phi_r(x + p\Delta,\ y + q\Delta)] \tag{7}$$

where $f_{pq}(x,y)$ is the p,q^{th} basis function generated from displaying the eigenvector

($\underline{\psi}_i^+ \Lambda_c^{-1} E_c^{-1} W_t^+$) in a two-dimensional format, $\phi_r(x,y)$ is a random function uniformly distributed from 0 to 2 , and Δ is a shift constant (see refs. 11-12 for more details of the coded-phase optical processor). Since the coded-phase function has an autocorrelation of a delta function, a coherent optical correlator produces an output consisting of the transform coefficients space at intervals of Δ

$$C(x', y') = \iint g(x,y) \exp\,[i\phi_r(x,y)] \sum_{p,q}^{M} f_{pq}{*}(x - x' + p\Delta,\ y - y' + q\Delta)$$

$$\exp[-\phi_r(x - x' + p\Delta,\ y - y' + q\Delta)]dxdy$$

$$= \sum_{p,q}^{M} \left[\iint g(x,y) f_{pq}{*}(x,y) dxdy \right] \delta\,(x' - p\Delta,\ y' - q\Delta) \tag{8}$$

Both the coded-phase distribution which illuminates the input and the spatial filter can be obtained by computer generated holograms.

Experimentally the optical Fukunaga-Koontz transformation was applied to the problem of distinguishing birds from fish. Ten images of song birds were input to the computer through a T.V. digitizer system. These images formed the training set for class 1. Ten images of fish were also input to the computer to form the training set for class 2. These two sets are shown in figure 2.

The F-K basis functions are shown in figure 3(a) along with the test images consisting of five new birds and five new fish in figure 3(b). Since the basis functions can contain negative values, the pictures have been scaled so that black equals the most negative, and white the most positive. The grey level which corresponds to a value of zero is shown in the small square below each basis function. The eigenvalues corresponding to the ten basis functions are given in table 1. We see that the best basis function with bird-type features is number 8. The best basis function with fish-type features is number 3.

It is interesting to compare the third and eighth basis functions with the arithmetic average of each training set. Figure 4 shows that the most important F-K basis function for a class is similar, but not identical to the average filter for that class. This is expected since both means were retained in the training sets. However, the F-K basis functions are not all positive. The grey level corresponding to zero is shown in the small square under each basis function.

A filter was generated which contained six of the basis functions shown in figure 3(a) in phase-coded form. Basis functions numbers 3, 4, and 1 were chosen to represent fish-like features, and 8, 9 and 10 to represent bird-like features. This filter was placed in the filter plane of an optical correlator, and a computer hologram of the coded-phase array was placed in front of the input plane so that its reconstruction illuminated the input. This is illustrated in figure 5. The output was detected with a T.V. camera, digitized, and displayed on a T.V. monitor. The digital computer can measure these six coefficients, and use them as input data to a linear or non-linear classifier for best classification.

The results of using three of the ten test images in the coded-phase processor are shown in figures 6(a)-(c). The lower right-hand corner of the T.V. monitor contains a sampled version of the actual light field detected by the T.V. camera at the output of the coded-phase processor. It consists of six points of light, where the square root of the intensity of each point corresponds to the absolute value of the coefficient for a specific basis function. The computer has added the lines around these points and the corresponding basis function numbers to help the display to be more meaningful. The measured values of the coefficients are plotted by the computer in the lower left-hand part of the screen. The input image has been reproduced in the upper left.

A simple linear classifier based on the coefficients from basis functions 3 and 8 was performed by the digital computer. This is drawn by the computer in the upper right-hand corner of the screen. The magnitudes of the two coefficients define a point in the two-dimensional space which is located by the mark "x". Points which are located below and to the right of the dotted diagonal line are classified as birds, whereas points above and to the left of the line are classified as fish. It is clear from the figures that in each case, the combination of an F-K transform and a simple linear classifier leads to a correct classification of the input image. Figure 7 shows the combined result of 30 classifications, based on the magnitude of coefficients from basis functions 3 and 8. Twenty of the points are from the training set of figure 2 and ten are from the test set of figure 3(b). A linear decision surface is able to separate the class "birds" from the class "fish" with no errors.

III. THE LEAST-SQUARES LINEAR MAPPING TECHNIQUE

The LSLMT is useful for performing a transform A, which maps an image vector $\underline{X}_j^{(i)}$ from the i^{th} class in the large N-dimensional feature space as close to one specific unit $\underline{V}_i$ as possible in the K-dimensional decision space such that the overall mean-square error incurred in the mapping is minimized (see figure 8)

$$A\underline{X}_j^{(i)} = \underline{V}_i + \underline{\varepsilon}_{ij} \quad (i = 1, 2, \ldots K;\ j = 1, 2, \ldots, M) \tag{9}$$

where $\underline{\varepsilon}_{ij}$ is the error vector for the i^{th} classifier and j^{th} image sample. The algorithm of LSLMT involves maximizing the interclass differences and minimizing the intraclass variations.

A. Procedures for Calculating the Basis Functions

The procedures for calculating the matrix A and the basis functions of the LSLMT are summarized as follows: (Details can be found in ref. 10.)

(a) Given a training set of images $\{\underline{X}_j^{(i)}\}$ which consists of K image classes and M sample images in each class (i = 1, ..., K; j = 1, ..., M), we define a (N × KM) matrix W as

$$W = \left[\underline{X}_1^{(1)}, \underline{X}_2^{(1)}, \ldots, \underline{X}_M^{(1)} \vdots \underline{X}_1^{(2)}, \underline{X}_2^{(2)}, \ldots, \underline{X}_M^{(2)} \vdots \ldots \vdots \underline{X}_1^{(K)}, \underline{X}_2^{(K)}, \ldots, \underline{X}_M^{(K)}\right] \tag{10}$$

(b) We choose the unit vectors $\underline{V}_i$ in the K-dimensional decision space to be orthonormal, e.g., when K = 3,

$$\underline{V}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \underline{V}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \underline{V}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix} \tag{11}$$

(c) Knowing the vectors $\underline{X}_j^{(i)}$ and $\underline{V}_i$ from parts (a) and (b) respectively, we can calculate $\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{V}_i \underline{X}_j^{(i)+}$.

(d) Knowing the matrix W from part (a), we can find the eigenvectors Ψ and eigenvalues Λ of W^+W

$$W^+W = \Psi\Lambda\Psi^+ \tag{12}$$

It is noted that W^+W is a $(KM \times KM)$ matrix.

(e) It can be shown that the inverse of the correlation matrix is given by

$$\left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{X}_j^{(i)} \underline{X}_j^{(i)+}\right]^{-1} = W\Psi\Lambda^{-2}\Psi^+W^+ \tag{13}$$

(f) Using the results of parts (c) and (e), we can calculate the matrix A for LSLMT.

$$A = \left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{V}_i \underline{X}_j^{(i)+}\right] \left[\sum_{i=1}^{K} \sum_{j=1}^{M} \underline{X}_j^{(i)} \underline{X}_j^{(i)+}\right]^{-1} \tag{14}$$

The basis functions are given by the row vectors of matrix A.

B. Optical Implementation of LSLMT

To employ the coded-phase processor for performing the matrix-vector multiplication of equation (9), we first need to convert the column vector $\underline{X}_j^{(i)}$ corresponding to the input image g(x,y) back to its two-dimensional image format of $(n \times n)$ pixels and the row vectors of the matrix A to the pattern functions $f_{pq}(x,y)$. The total number of pattern functions is K, i.e., $p+q = 1, 2, \ldots, K$, because there are only K row vectors in the matrix A. The output from the coded-phase processor will then be in a two-dimensional form with a spacing Δ between different light spots whose magnitudes of brightness are V_i, $i = 1, 2, \ldots, K$. The brightest spot indicates which class the input image belongs to.

The key element required in optically implementing LSLMT is a computer generated hologram filter whose impulse response contains the pattern functions f_{pq} derived from the row vectors of matrix A. The advantages of optical classification of statistical patterns are parallel processing and real-time data rates. As an example, to recognize any unknown image of $(n \times n)$ pixels which has the same statistical properties as those images used in the training sets of K-classes, the amount of parallel computations involved is $(K \times n \times n)$ multiplications and $(K \times n \times n)$ additions in real-time. When the real-time rate is 1/30 sec, K is 100, n is 500 and each pixel has 64 gray levels (or 6 bits), the computational rate is greater than 9×10^9 bits/sec.

Experimentally, the LSLMT was applied to design a classifier for hand-written letters in a variety of styles. Ten images of each of the hand-written characters m, t and a (as shown in figure 9) were input to the computer through a TV camera/digitizer system forming the training sets of three classes (detail description about our micro-computer based video-image analysis system are given in refs. 11-12). Using the thirty image samples for the three classes and following the procedures outlined in Section IIIA, we calculated the matrix A for LSLMT. The three row vectors in A yield the three pattern functions f_{pq} as shown in figure 10. Using these pattern functions we next generate a computer hologram filter whose impulse response is given by equation (7). The computer generated filter is shown in figure 11 and applied to the eight test images shown in figure 12.

Three of the eight results from optical implementation of LSLMT are illustrated in figure 13. On the right the outputs from the coded-phase processor are displayed, given the test patterns on the left. Each output contains three spots of light in pre-determined locations. The square roots of the spot intensities correspond to the absolute values of V_1, V_2 and V_3. The location of the brightest spot in the output plane clearly indicates classification. The measured values of V_i for all eight test patterns are summarily listed in table 2 and plotted in figure 14. Based on table 2 and figure 14, we can conclude that we have successfully demonstrated the optical implementation of LSLMT.

IV. COMPARISONS AMONG VARIOUS OPTICAL PATTERN RECOGNITION METHODS

We now turn to compare the LSLMT with other methods of optical pattern recognition and classification. The comparisons cover the matched filter, the average filter and the Fukunaga-Koontz transform. The first two comparisons were carried out digitally. The third comparison was optical.

A. Comparison with the Matched Filter

Two digital filters were generated with the impulse responses of the first letters m or t in the training sets shown in figure 9(a),(b). All the letters of m and t in the training sets (fig. 9(a),(b)) and the test set (fig. 15(a)) were used as the inputs g(x,y). Figure 15(b) shows the normalized correlations between the inputs and the matched filters, which are defined by

$$V_1 = \langle g m_1^* \rangle / \langle m_1 m_1^* \rangle$$

$$V_2 = \langle g t_1^* \rangle / \langle t_1 t_1^* \rangle \qquad (15)$$

where <> signifies an inner product. Using the dotted line as the classification boundary will obviously yield zero error of classification. However, a more restrictive classifier defined as the ratio between V_1 and V_2 is often preferred because it is insensitive to scalar multiplication of g(x,y) which may result from changes in input illumination. This more restrictive classification boundary passes through the origin, as shown by the solid line in figure 15, and yields 30% error in recognition.

This digital matched filtering result is compared with LSLMT using the same set of letters in figures 9(a),(b), and 15(a). Figure 16 shows that the LSLMT can provide error free classification.

B. Comparison with an Average Filter

Digital average filters were calculated by averaging the characters of the training sets in figure 9. The digital filters are shown in figure 17, which can be compared with the pattern functions f_{pq} for LSLMT (fig. 10). While the average filters are formed by superposing only the images of the same class, to form pattern functions f_{pq} for LSLMT, it appears that one may superpose not only the images of one class but also subtract the superposed images from other classes.

Using the digital average filters, we obtain the correlations V_1 and V_2 between the averaged characters and the input g(x,y), which can be any image in the training (fig. 9(a),(b)) or test (fig. 15(b)) sets.

$$V_1 = \langle g\sum_j m_j^* \rangle / \langle |\sum_j m_j|^2 \rangle$$

$$V_2 = \langle g\sum_j t_j^* \rangle / \langle |\sum_j t_j|^2 \rangle \tag{16}$$

Figure 18 shows the results of using digital average filters. The error rate is 20%, which is a little better than the 30% rate associated with digital matched filters. But, both digital average and matched filters yield results inferior to the LSLMT, which has a zero error rate.

C. Comparison with Fukunaga-Koontz Transform

This comparison was performed optically using the coded-phase optical processor and computer generated hologram filters. Two computer hologram filters were generated using the same training sets of images of song birds and fish, one based on the principles of the FKT described in reference 9 and the other based on the LSLMT (ref. 10). The results of optically implementing the LSLMT are shown in figure 19. Two of the six test images used as input are shown in the left column. The corresponding outputs each containing two spots of light for this two-class classification problem are shown in the middle column. The linear classification results are shown in the right column. These linear classification results together with those for four other test images are shown in figure 20(a), which are to be compared with the linear classification results based on the FKT as shown in figure 20(b). It is obvious that both FKT and LSLMT are good classifiers. The FKT has the advantage that it can provide many features for each image. These features can be used as the basis for a nonlinear classification routine. However, it is useful only for two-class problems. The LSLMT provides only K features (one feature for each class), but it can be applied to multiclass problems (K > 2).

V. SUMMARY

We have shown that the Fukunaga-Koontz transform can be used as a feature extractor in a two-class classification application. It was used for classifying birds and fish. After the F-K basis functions were calculated, those most useful for classification were incorporated into a computer generated hologram. A coherent optical processor was designed using this computer generated hologram to perform the F-K transform in real time. The output of the optical processor, consisting of the squared magnitude of the F-K coefficients, was detected by a T.V. camera, digitized, and fed into a micro-computer for classification. A simple linear classifier based on only two F-K coefficients was able to separate the images into two classes, indicating that the F-K transform had chosen good features.

The Least-Squares Linear Mapping Technique has been optically implemented to classify large images also by utilizing the phase-coded optical processor and computer generated hologram filters. The principles of LSLMT are incorporated into the computer hologram filters. Although we have demonstrated here only two- and three-class problems, the LSLMT is useful for optical classification problems of many classes.

A method was developed which simplified the computation of the FKT or LSLMT basis functions for large dimensional imagery, and was found to work well when the intraclass variation in each class was small, and the correlation matrix could be approximated by a sample correlation matrix of low rank.

The advantages of optically implementing the FKT and LSLMT are parallel and real-time processing. In comparisons with the matched filter and the average filter, the LSLMT is clearly superior for classifying statistical patterns because it maximizes the interclass differences and minimizes the intraclass variations.

ACKNOWLEDGEMENTS

The research was performed in collaboration with J. Leger and Z. Gu. The support of this research from the National Science Foundation and the Air Force Office of Scientific Research is gratefully acknowledged.

REFERENCES

1a. Vander Lugt, A.: Signal Detection by Complex Spatial Filtering. IEEE Trans. Inform. Theory, vol. IT-10, 1964, pp. 139-145.

1b. Vander Lugt, A., F. B. Rotz and A. Klooster: Character Reading by Optical Spatial Filtering. In Optical and Electrooptical Information Processing, Tippet et al., eds., MIT Press, Cambridge, Mass., 1965, p. 125.

2. Gara, A. D.: Real-time Tracking of Moving Objects by Optical Correlation. Appl. Opt., vol. 18, 1979, pp. 172-174.

3. Almeida, S. P. and J. K. T. Eu: Water Pollution Monitoring Using Matched Spatial Filters. Appl. Opt, vol. 15, 1976, pp. 510-515.

4. Casasent, D. and D. Psaltis: Position, Rotation, and Scale Invariant Optical Correlation. Appl. Opt., vol. 15, 1976, pp. 1795-1799.

5. Hester, C. F. and D. Casasent: Optical Pattern Recognition Using Average Filters to Produce Discriminant Hypersurfaces. Proc. SPIE, vol. 201, 1979, pp. 77-82.

6. Caulfield, H. J., R. Haimes and D. Casasent: Beyond Matched Filtering. Opt. Eng., vol. 19, 1980, pp. 152-156.

7. Leger, J. R. and S. H. Lee: A Hybrid Optical Processor for Pattern Recognition and Classification Using a Generalized Set of Pattern Functions. Appl. Opt., 1982 (to be published).

8. Duvernoy, J: Optical Pattern Recognition and Clustering: Karhunen Loéve Analysis, Appl. Opt., vol. 15, 1976, pp. 1584-1590.

9. Leger, J. R. and S. H. Lee: Optical Implementation of Fukunaga-Koontz Transform. J. Opt. Soc. Amer., 1982 (to be published).

10. Gu, Z. H., J. R. Leger and S. H. Lee: Optical Implementation of the Least-Squares Linear Mapping Technique for Image Classification. J. Opt. Soc. Amer., 1982 (to be published).

11. Leger, J. R. and S. H. Lee: Coherent Optical Implementation of Generalized Two-Dimensional Transforms. Opt. Eng., vol. 18, 1979, pp. 518-523.

12. Leger, J. R., J. Cederquist and S. H. Lee: Micro-computer Based Hybrid Processors at UCSD. Opt. Eng., 1982 (to be published).

Table 1. Eigenvalues Corresponding to Ten Basis Functions of FKT

Basis Function	Eigenvalue (λ_i)	$\lvert\lambda_i - 0.5\rvert$	Best for which class?
1	0.1354	0.3646	fish
2	0.4846	0.0154	fish
3	0.0008	0.4992	fish
4	0.0034	0.4966	fish
5	0.9416	0.4416	birds
6	0.9883	0.4883	birds
7	0.9966	0.4966	birds
8	0.9998	0.4998	birds
9	0.9992	0.4992	birds
10	0.9985	0.4985	birds

Table 2. The measured values of V_1, V_2 and V_3 for the eight test patterns shown in Figure 12.

	V_1	V_2	V_3	Largest for which class?
m_1	1.58	0.02	0.12	m
m_2	0.93	0	0.12	m
t_1	0.02	1.1	0.07	t
t_2	0.07	1.52	0.01	t
t_3	0.05	0.95	0.02	t
a_1	0.02	0.01	0.88	a
a_2	0.03	0.01	1.63	a
a_3	0.02	0.01	1.22	a

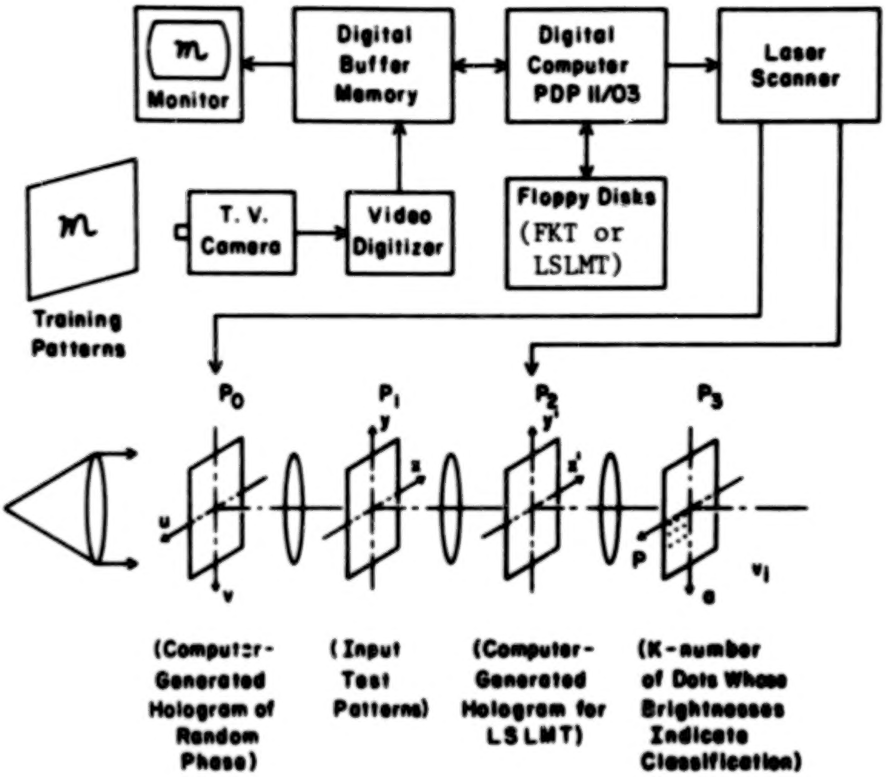
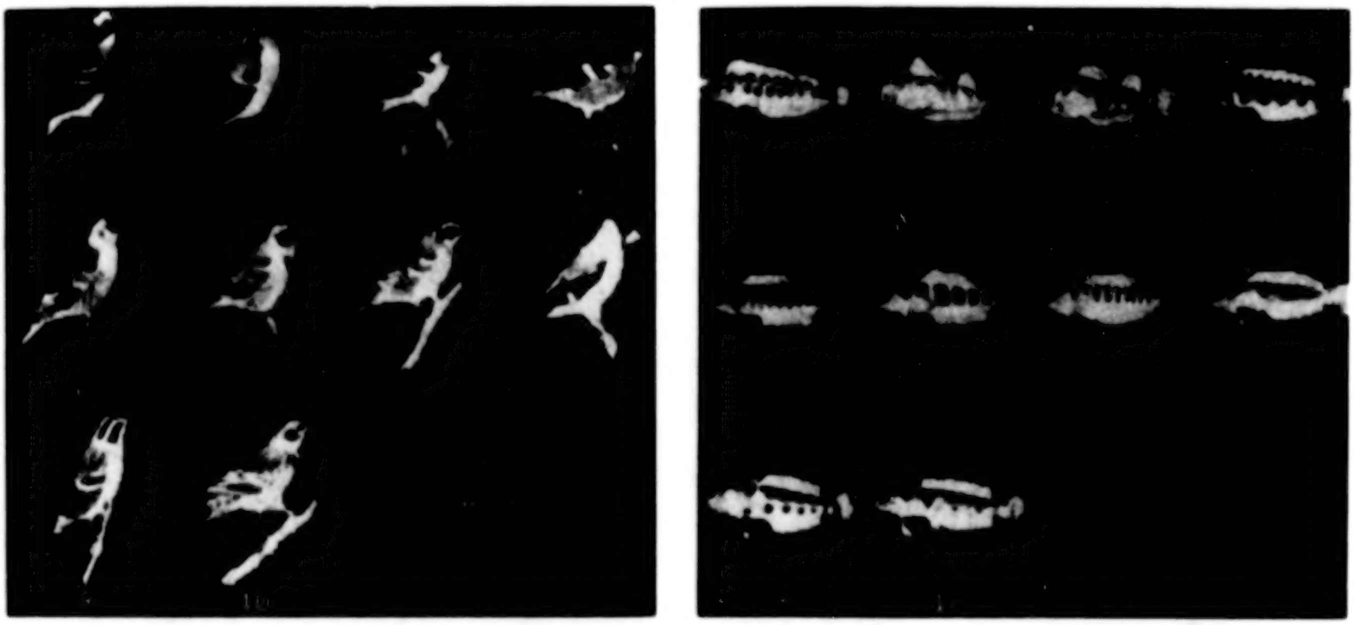
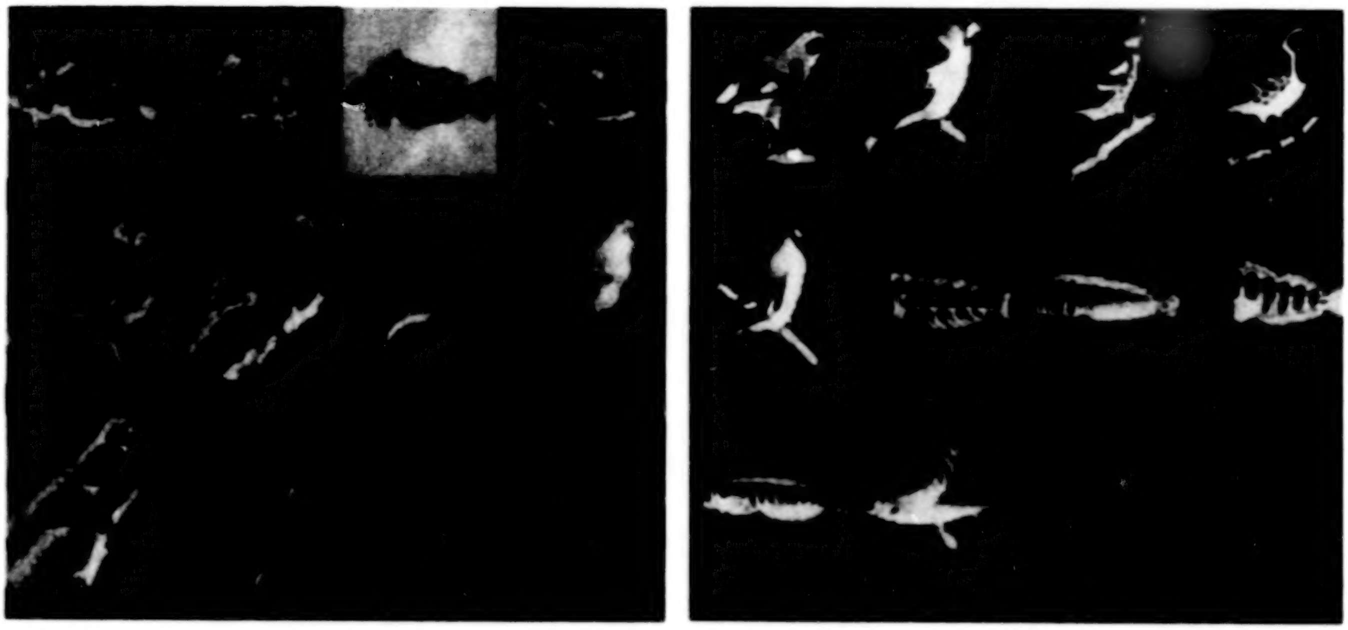

Monitor
Digital Buffer Memory
Digital Computer PDP 11/03
Laser Scanner
Training Patterns
T. V. Camera
Video Digitizer
Floppy Disks (FKT or LSLMT)
P0
P1
P2
P3
u
v
y
x
y'
x'
p
q
v1
(Computer-Generated Hologram of Random Phase)
(Input Test Patterns)
(Computer-Generated Hologram for LSLMT)
(K-number of Dots Whose Brightnesses Indicate Classification)

Figure 1.- The hybrid system for optical recognition of statistical patterns.



(a) (b)

Figure 2.- (a) Class 1 training set consisting of ten song birds. (b) Class 2 training set consisting of ten fish.



(a) (b)

Figure 3.- (a) F-K basis functions. Basis function 3 is best for class 2 (fish), and basis function 8 is best for class 1 (birds). Small square to the left of image number indicates grey level corresponding to zero. (b) Test images consisting of five new birds and five new fish.

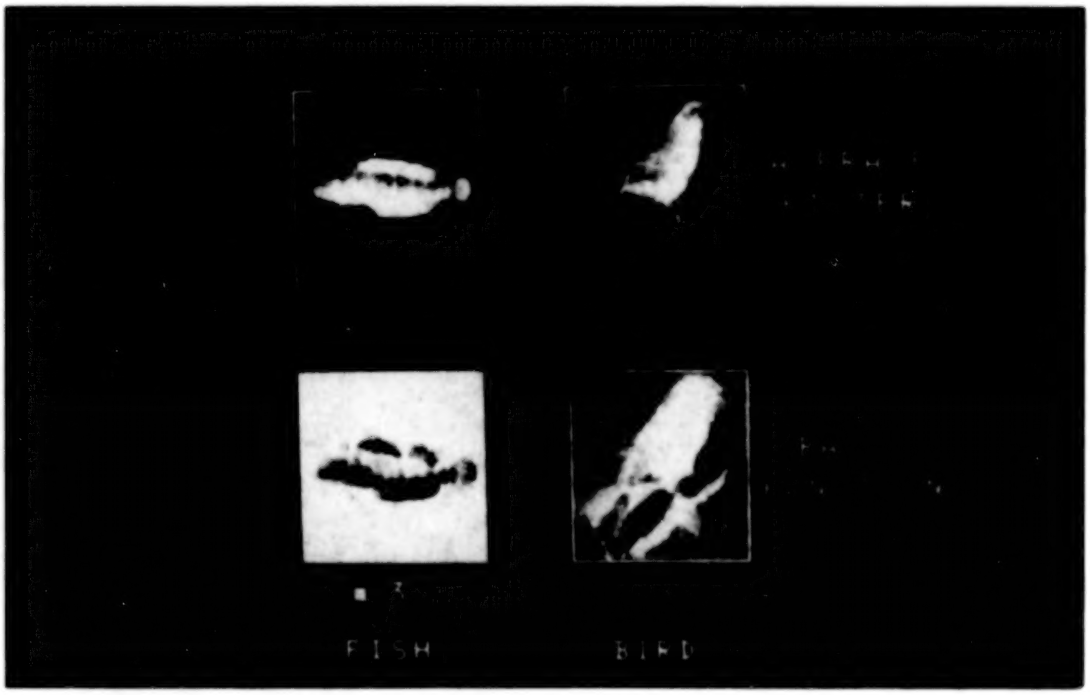


Figure 4.- Comparisons of basis functions 3 (fish) and 8 (birds) with filters formed by the arithmetic average of the training sets.

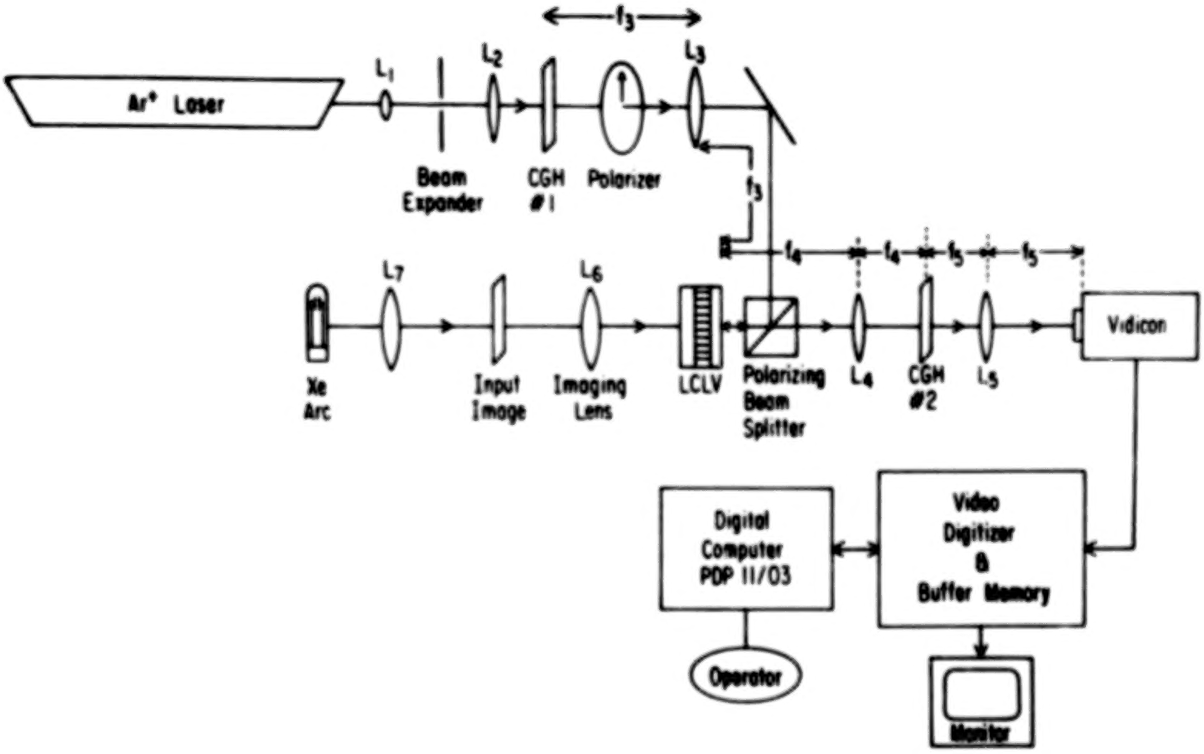
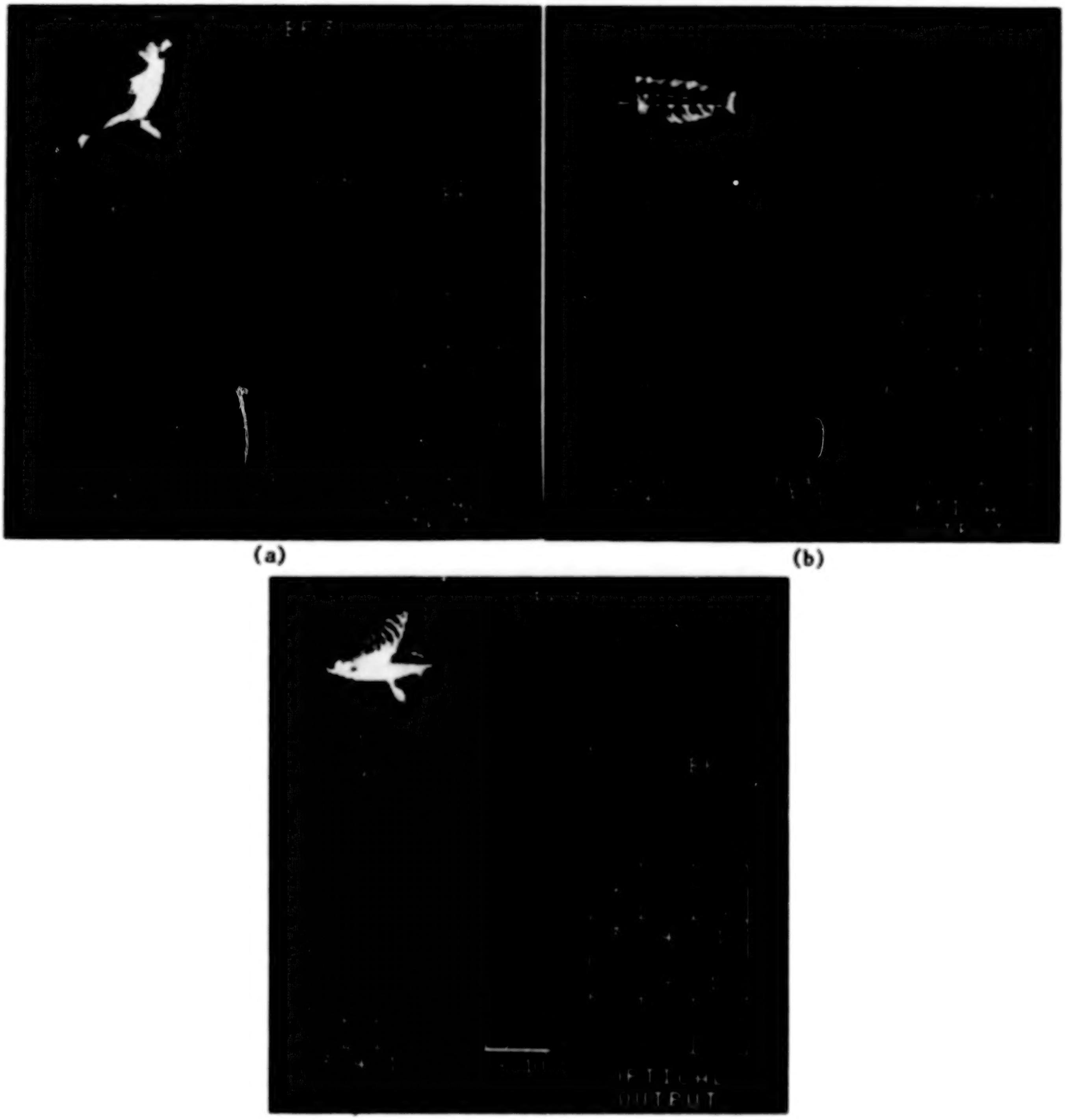

f3
Ar+ Laser
L1
L2
L3
Beam Expander
CGH #1
Polarizer
f3
f4
f4
f5
f5
L7
L6
Xe Arc
Input Image
Imaging Lens
LCLV
Polarizing Beam Splitter
L4
CGH #2
L5
Vidicon
Digital Computer PDP 11/03
Video Digitizer & Buffer Memory
Operator
Monitor

Figure 5.- Hybrid implementation of the coded-phase optical processor. A computer hologram of a coded-phase array is shown as CGH#1. A second computer hologram containing the F-K basis functions in coded-phase form is shown as CGH#2. LCLV is a liquid crystal light valve for converting an incoherent (test) image into a coherent image. The resultant F-K coefficients are detected by the vidicon and analyzed by a digital computer. With the same CGH#2 but new test images, new F-K coefficients are obtained and new classifications are achieved in real time.



(a) (b)

(c)

Figure 6.- Optical implementation of the F-K transform using the coded-phase optical processor. The six basis functions with greatest separation power were used. Basis functions 3, 4 and 1 are best for fish, and 8, 9 and 10 are best for birds. Three of the ten images are used as inputs in (a), (b) & (c), where the input is reproduced in the upper left-hand corner of the T.V. screen. The output of the optical processor is detected by a T.V. camera, displayed in the lower right-hand corner of the screen, measured by the video digitizer, and graphed in the lower left. A linear classifier using basis functions 3 and 8 is shown in the upper right-hand corner of the screen, with the dotted line separating birds (below and to the right of the line) from fish (above and to the left of the line).

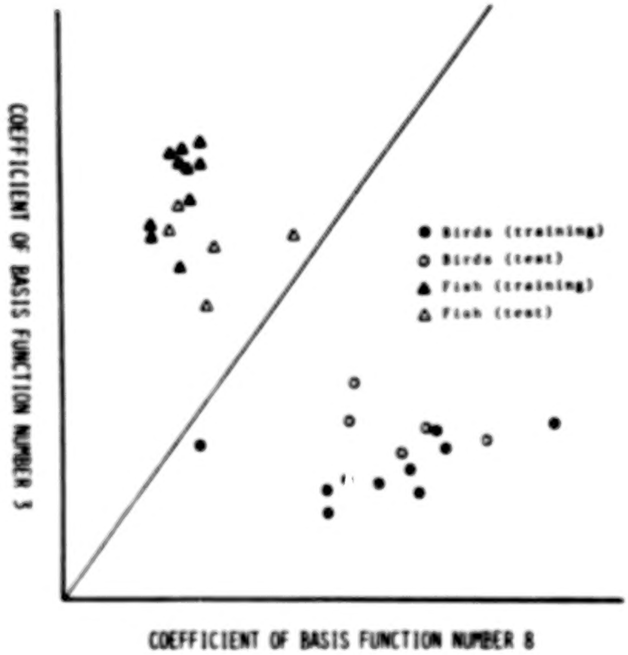


Figure 7.- Classification of birds and fish using coefficients of basis functions 3 and 8. Ten members of each training set, as well as ten members of the test set were classified.

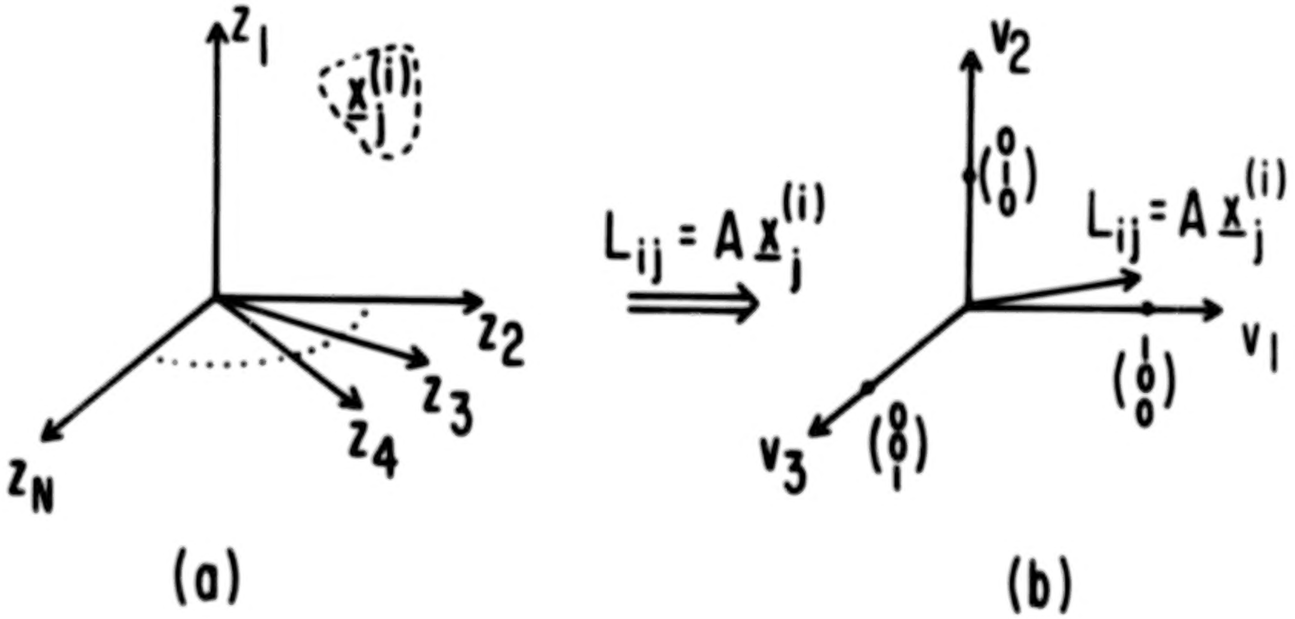

z_1
$\underline{x}_j^{(i)}$
z_2
z_3
z_4
z_N
$L_{ij} = A\,\underline{x}_j^{(i)}$
v_2
$\begin{pmatrix}0\\1\\0\end{pmatrix}$
$L_{ij} = A\,\underline{x}_j^{(i)}$
v_1
$\begin{pmatrix}1\\0\\0\end{pmatrix}$
v_3
$\begin{pmatrix}0\\0\\1\end{pmatrix}$
(a)
(b)

Figure 8.- The schematic diagrams of LSLMT for 3 classes. (a) N-dimensional feature space. (b) Three-dimensional decision space.

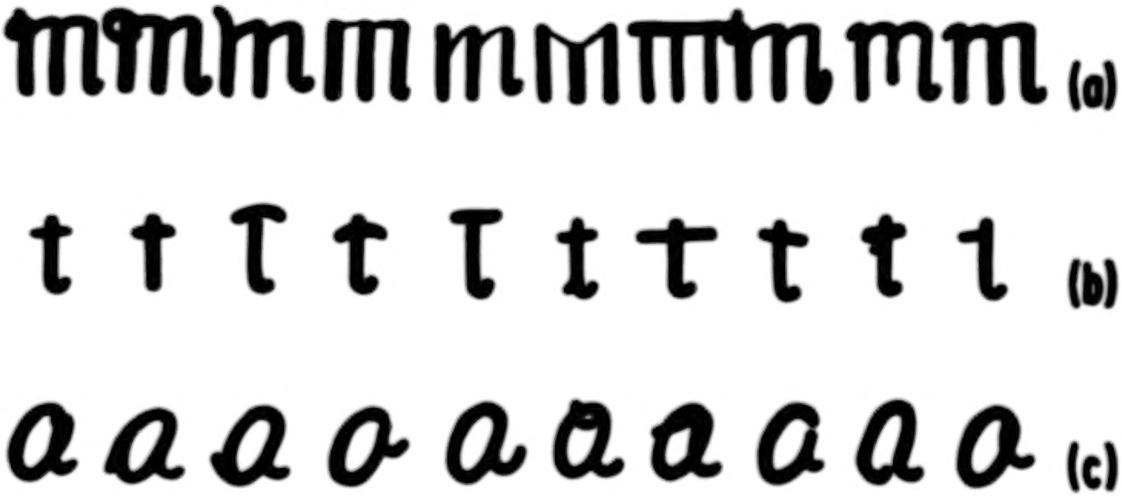


Figure 9.- (a) Class 1 training set consisting of ten characters of m. (b) Class 2 training set consisting of ten characters of t. (c) Class 3 training set consisting of ten characters of a.

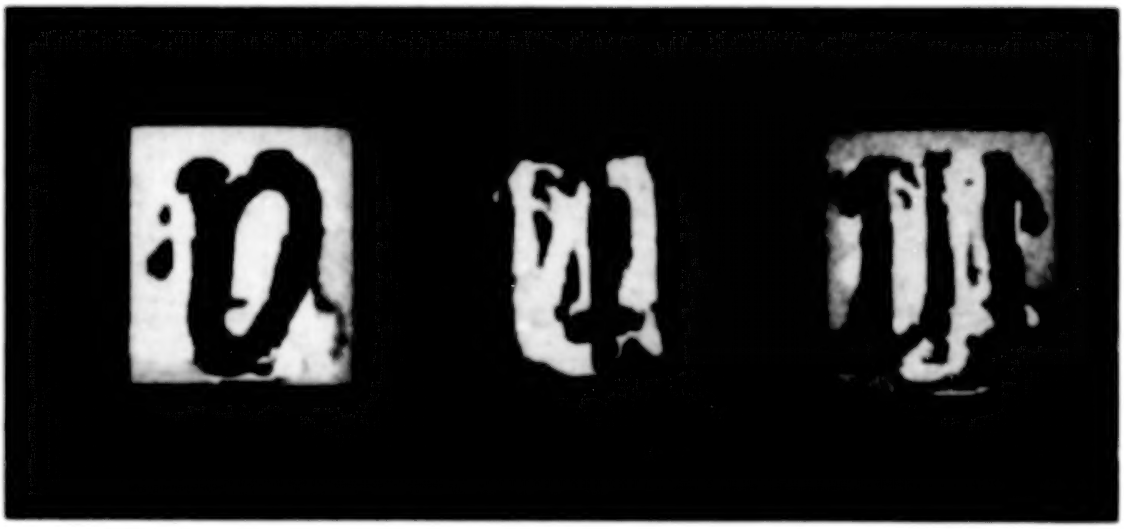
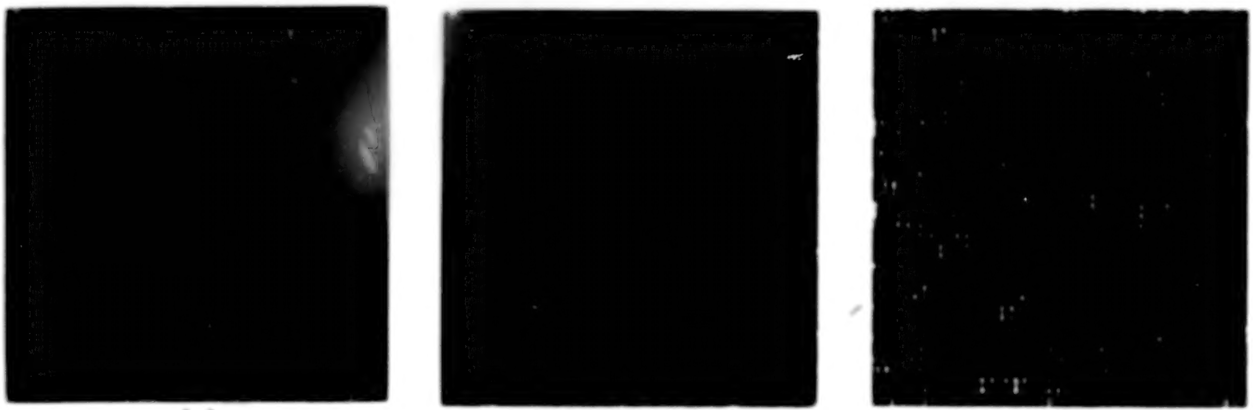


Figure 10.- Three pattern functions f_{pq} for LSIMT. Since the pattern functions f_{pq} contain both positive and negative values in general, bias levels (indicated by the small grey squares below f_{pq}) are added to f_{pq} to display them.



(a) (b) (c)

Figure 11.- A computer-generated spatial filter for LSLMT made with a laser scanning system. (a) 4 × 4 filter array, (b) one filter, (c) the central part of (b).

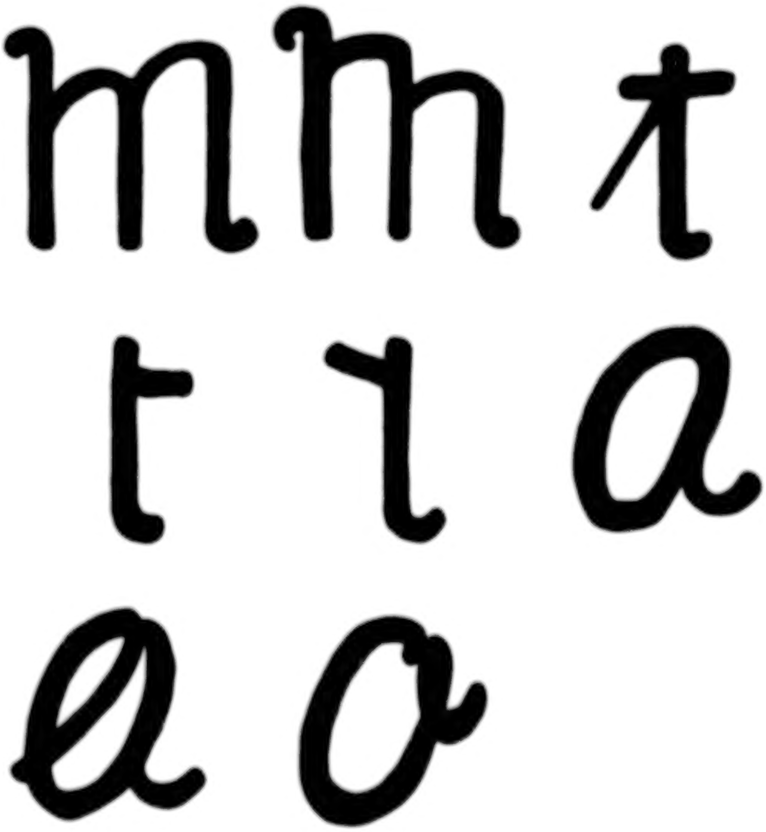


Figure 12.- A test set of images consisting of two new m-characters, three new t-characters and three new a-characters.

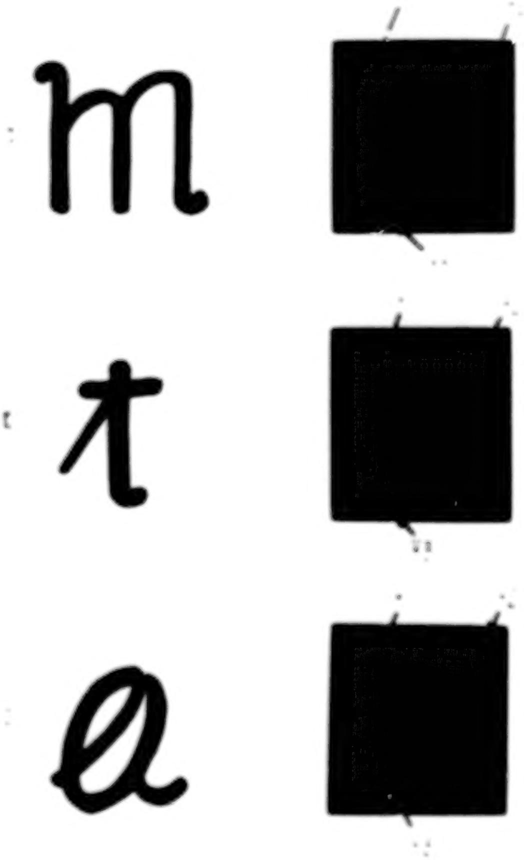


Figure 13.- The result of optical implementation of the LSLMT for 3 classes. Three of the eight test images used are shown on the left. The outputs of the optical processor are displayed on the right.

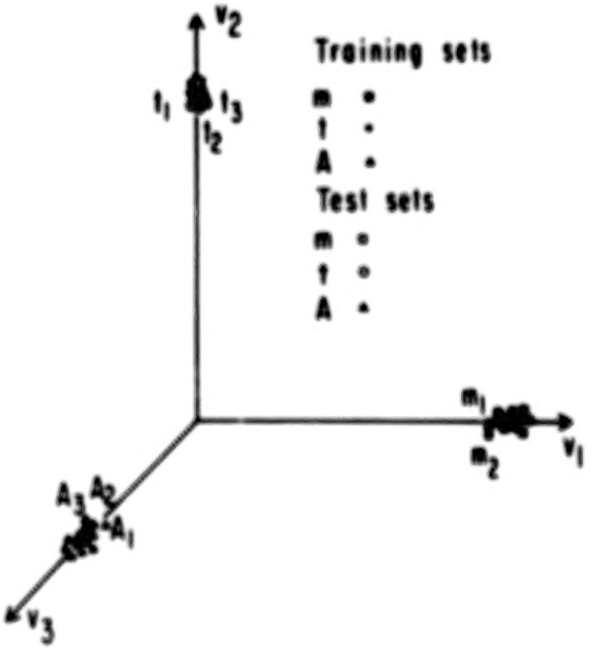


Figure 14.- Classification result with LSLMT for three classes (m,t,a). The solid symbols of triangle, circle and square are for the training sets. The hollow symbols of triangle, circle and square are for the test images.

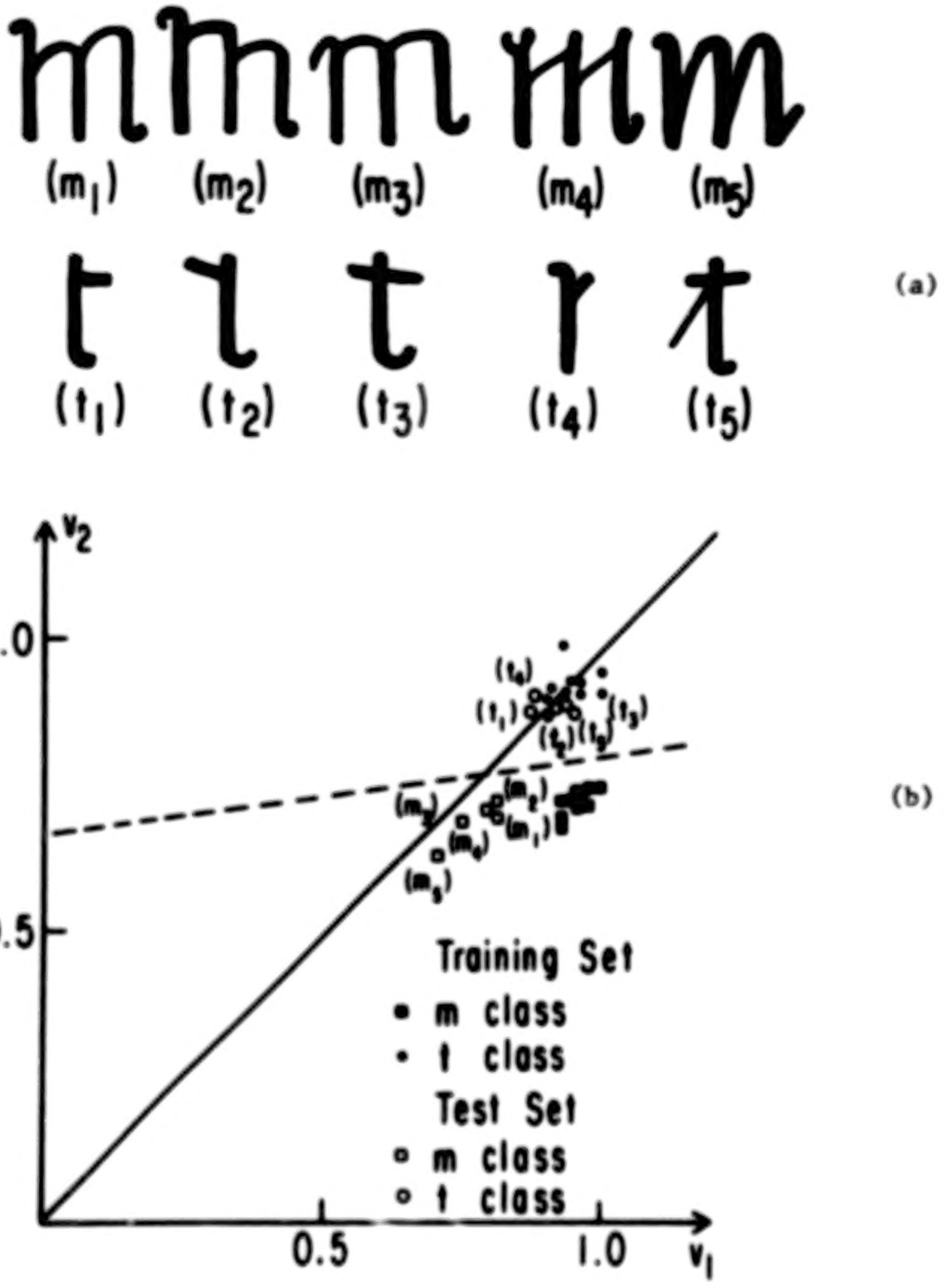


Figure 15.- (a) A test class of m and t characters. (b) Classification using the matched filter for 2 classes (m,t). The solid line represents a restrictive linear classification boundary. The error rate is 9 out of 30, or 30%.

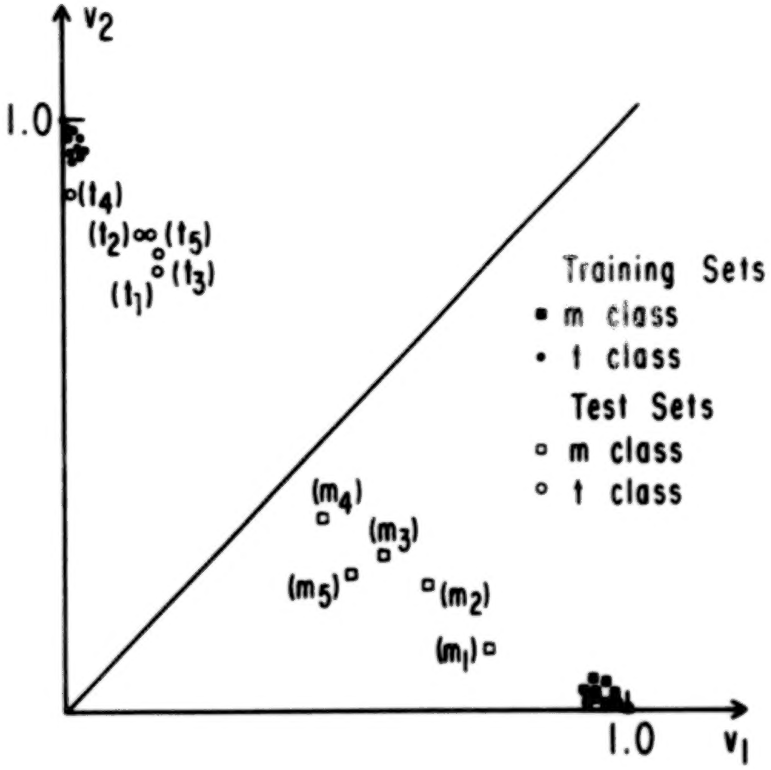


Figure 16.- Classification result with LSLMT of 2 classes (m and t characters).

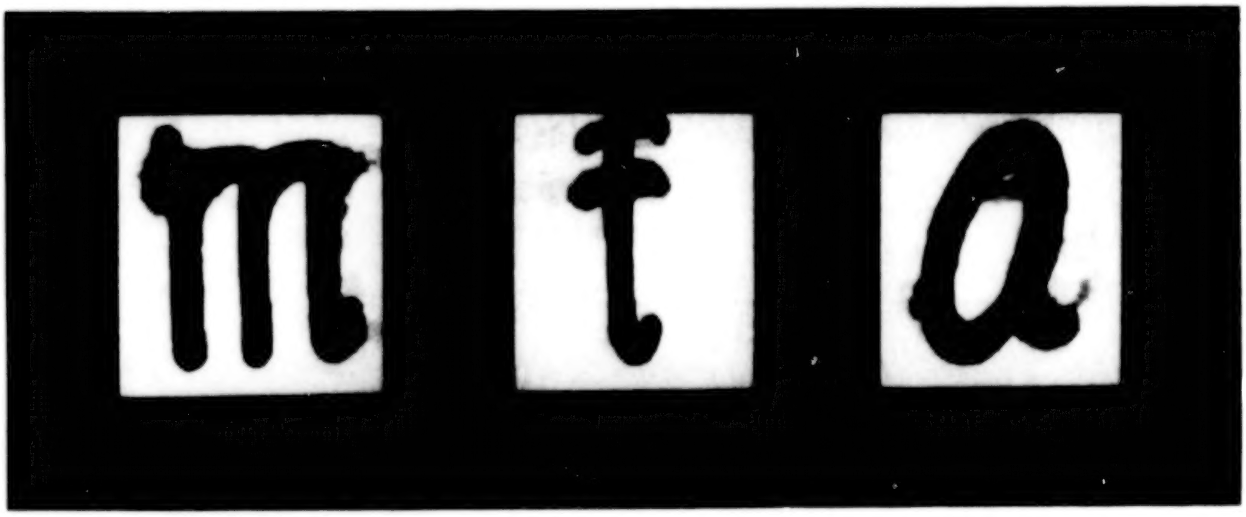


Figure 17.- 3 average filters of the training sets (m,t,a).

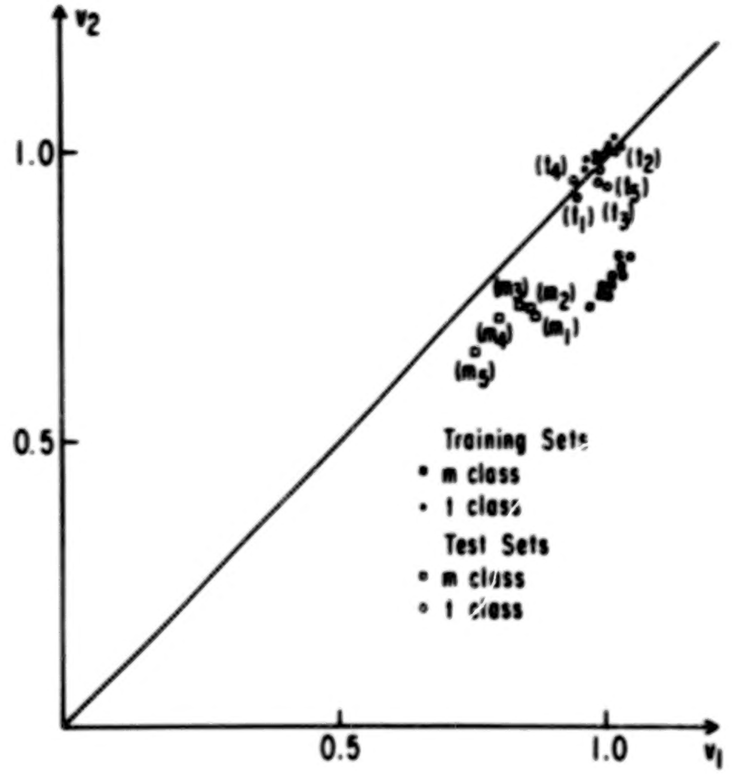


Figure 18.- Classification using average filters for 2 classes (m,t). The error rate is 6 out of 30, or 20%.

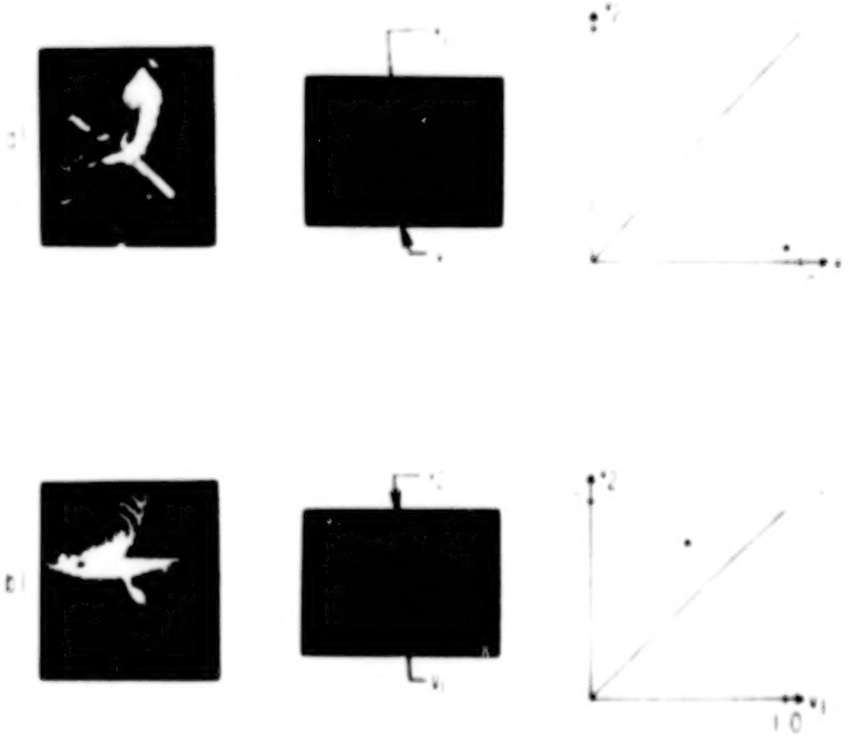


Figure 19.- The optical implementation of the LSLMT for two classes (Song Bird and Fish). The left column shows two of the six test images as input, the middle column shows the outputs, the linear classifier is shown in the right column.

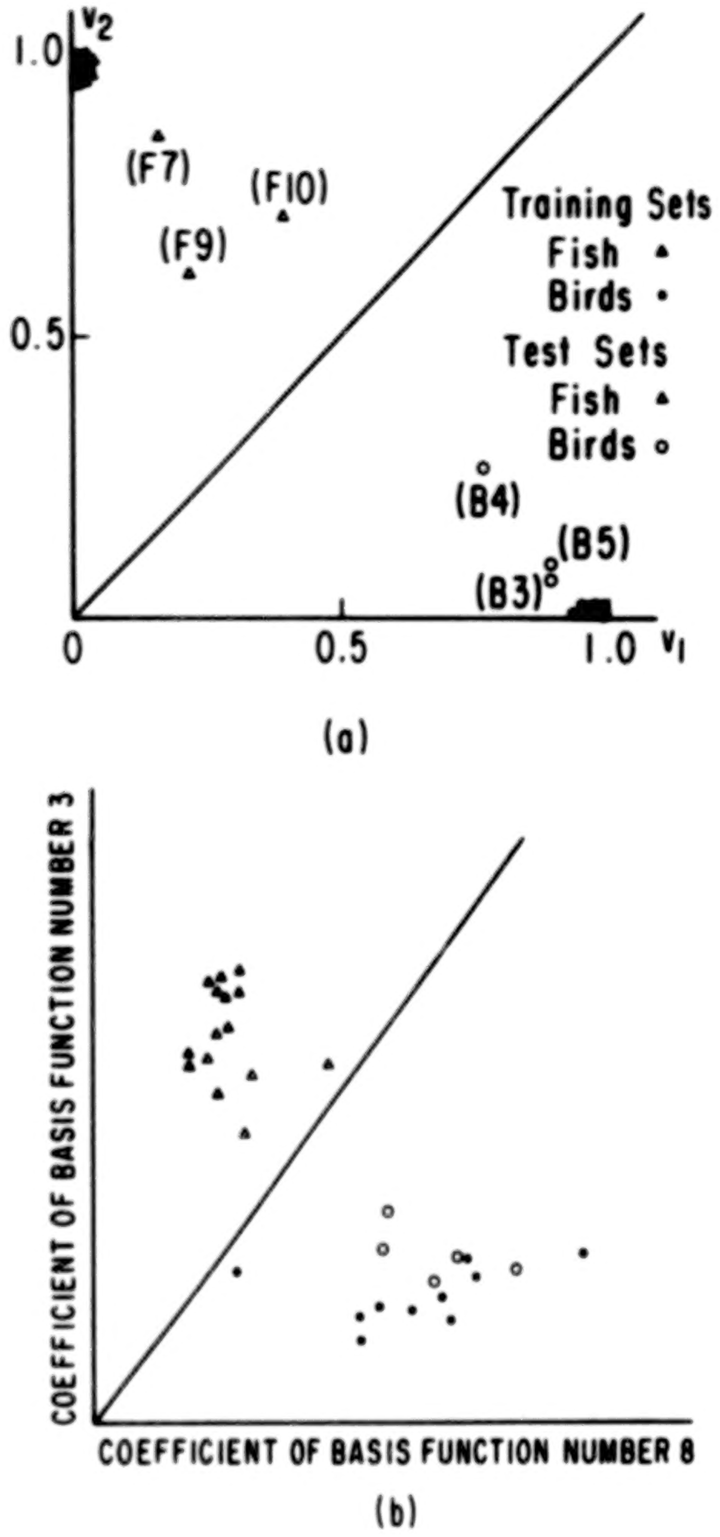


Figure 20.- (a) Results of optically implementing the LSLMT for two classes (Song Bird and Fish). (b) Results of optically implementing the Fukunage-Koontz transform using coefficients of basis functions 3 and 8.

OPTICAL SPATIAL INTEGRATION METHODS FOR AMBIGUITY FUNCTION GENERATION

P.N. Tamura, J.J. Rebholz, O.T. Daehlin, and T.C. Lee

Honeywell Corporate Technology Center

Bloomington, MN. 55420

SUMMARY

An optical spatial integration approach to ambiguity function generation has been developed. It uses acousto-optic Bragg cells as input transducers and a special passive optical element called a space variant linear phase shifter to create time-delayed versions of one of the signals. Real time processing has been achieved.

INTRODUCTION

The ambiguity function $\chi(\nu,\tau)$ for two given signals $f_1(t)$ and $f_2(t)$ is defined by:

$$\chi(\nu,\tau) = \int_{-\infty}^{\infty} f_1(t)\, f_2^*(t-\tau)\, \exp(-j2\pi\nu t)\, dt. \tag{1}$$

The ambiguity function is widely used in the processing of radar and sonar signals. A received signal $f_2(t)$ will differ from a reference signal $f_1(t)$ by a time delay and a Doppler shift corresponding to the range and radial velocity component of a target. When the ν and τ variables in (1) take on the appropriate values to compensate for the delay and Doppler differences between $f_1(t)$ and $f_2(t)$, the ambiguity function will yield a peak value at the specific position (ν,τ) in the range-Doppler plane.

Equation (1) imposes a severe computational load on digital signal processing hardware, especially in surveillance applications where timely computation of large numbers of ambiguity functions calls for near real time processing. Consequently, many optical schemes to perform the ambiguity calculation have been investigated over the years. These include both spatial integration approaches[1-6], and more recently, time integrating architectures that utilize acousto-optic Bragg cells as input transducers.[7]

In this paper, a coherent spatial integration optical processor architecture that computes the ambiguity function will be described. The results of real time processing will also be shown.

PROCESSING ARCHITECTURE

Inspection of equation (1) shows that the operations to be performed by the sign is opposite.

Therefore, the space variant linear phase shifter can be accurately fabricated by cementing a cylindrical lens and a spherical lens of opposite power together, and orienting them at 45°. The focal length of the cylindrical lens should be half that of the spherical lens.

ACOUSTO-OPTIC BRAGG CELLS

The key element in any optical processing system is the input transducer which converts an electronic signal in time to an optical signal in space. The fastest such device known is the acousto-optic Bragg cell, which receives a temporal signal modulated on an RF carrier which propagates through the crystal as a shear wave, and space-modulates an incident optical beam. Speed is a critical factor in ambiguity function calculation, and since digital signal processor throughput will soon approach the video rates at which most two-dimensional spatial light modulators operate, Bragg cells, which have a frame time of 50 μs, are a very appropriate choice as input transducers. In addition to having a fast frame time, Bragg cells also have the advantages of electronic data composing and high diffraction efficiency.

The Bragg cells used for all experiments were TeO_2 devices that operate in the slow shear mode with 50 μs frame times. The center frequency was 56.[illegible] MHz with 40 MHz bandwidth at the 3 dB points, giving a time-bandwidth product of 2000. They were designed to operate at 514.5 nm.

REAL TIME PROCESSING

The most important design criterion in the real time system is the system time-bandwidth product (TBW). An acoustic shear wave in a TeO_2 Bragg cell will propagate with a speed of 617 meters/sec, while the practical physical window size is 30 mm. It takes only 48 μsec for the acoustic wave to propagate through the window. A sufficient system TBW, therefore, requires a fast data rate and high speed DACs. Fast 8 bit DACs commercially available operate at a 50 MHz data rate, and a digital time base compression (TBC) buffer of comparable speed must be constructed to operate with it.

The real time system is designed around the optical ambiguity function generator that utilizes two Bragg cells and the LIPS element. The front end electronics consisting of a TBC buffer, two digital-to-analog converters (DACs), and two RF driver/modulators, receives the input signals from the HP 9825 calculator in digital form and converts them to analog signals which cause index of refraction modulations in the Bragg cells which are proportional to the input signals. The rear end electronics, consisting of a two-dimensional detector array camera and a video memory, captures the ambiguity function, displays it on a CRT, and digitizes and stores it for examination. A block diagram of the real time system is seen in Figure 3.

The processing rate of the real time system is limited by the detector array read out time to 500 frames/sec. Also, the HP 9825 cannot reload the TBC buffer fast enough to create 500 different frames each second. Instead, the same frame is optical signal processor are a conjugation operation, a multiplication of $f_1(t)$ by $f_2(t-\tau)$, and a one dimensional Fourier transform along the Doppler axis. These operations can easily be performed in the space domain by Fourier optics techniques, and an optical architecture capable of performing them is shown in Figure 1.

The first operation, conjugation, is performed by placing each signal on a carrier frequency and passing the opposite diffracted order from each Bragg cell. This heterodyning with a carrier is necessary anyway since the signal data is invariably in complex form from a previous basebanding operation. The second operation, multiplication, is performed by imaging Bragg cell I telecentrically on to Bragg cell II. Finally, lens S5 takes a one-dimensional Fourier transform along the frequency axis to yield the ambiguity function. This architecture has been analyzed in formal Fourier mathematics in a previous paper[8].

The time-shifted version of $f_2(t)$ is created by using a passive optical element called a space variant linear phase shifter (LPS). The action of this LPS element can be understood by considering that its horizontal cross section is that of a prism. The result of imaging Bragg cell I onto Bragg cell II through a prism is to cause a position shift of the image. Since this is a spatial integration approach, a misregistration in the space domain is equivalent to a time shift in the time domain, and a delayed version of $f_2(t)$ will be created. By continually varying the wedge angle of the prism along the vertical axis, a continuous range of delay values can be created.

The feasibility of implementing the optical architecture of figure 1 depends heavily on the manufacturability of the linear phase shifter element. It is essentially an optical wedge; its wedge angle changes linearly with height. The complex transmissivity function of this component is given in rectangular coordinates by

$$g(x,y) = e^{j\alpha xy} \tag{2}$$

where α is a constant.

Conventional manufacturing processes such as grinding and polishing a glass piece would be difficult if not impossible to apply to the fabrication of such an element. We have invented a method to fabricate this component out of conventional optics, hence high accuracy of the transmitted wavefront is possible.

By modifying equation (2), we have

$$g(x,y) = e^{j\frac{\alpha}{2}(x+y)^2} \; e^{-j\frac{\alpha}{2}(x^2+y^2)} \tag{3}$$

where the decomposition of equation (3) is essentially the same as that used in the chirp-z algorithm. If we define $r = x^2 + y^2$ and introduce a coordinate system (x',y') that is rotated from (x,y) by 45° (figure 2), equation (3) can be rewritten as

$$g(x,y) = e^{j\alpha x'^2} \; e^{-j\frac{\alpha}{2} r^2} . \tag{4}$$

The first exponent in equation (4) is the complex transmissivity function of a cylindrical lens oriented parallel to the x' axis. The second exponent is a spherical lens. The cylindrical lens is twice as powerful as the spherical lens, and the repetitively read to the Bragg cells to demonstrate high speed processing.

The test signals used to demonstrate the real time processing capability of this system were synthetically produced complex random signals with varying amounts of noise added. The time bandwidth product was selected to be 256. The output of the 100 x 100 detector array camera was plotted in an isometric display format. One such plot for an input signal-to-noise ratio of 0 dB (signal power equal to noise power) is shown in figure 4. The optically produced results compare favorably with digital calculations, and their resolutions along both the delay and Doppler axes are in agreement with theoretical predictions.

CONCLUSIONS

A coherent spatial integration approach to ambiguity function generation has been described. It uses one-dimensional acousto-optic Bragg cells in conjunction with a space variant linear phase shifter optical element to generate the two-dimensional ambiguity function in one exposure. A real time implementation of this system has been demonstrated.

REFERENCES

1. L.J. Cutrona, E.N. Leith, C.J. Palermo and L.J. Porcello, "Optical data processing and filtering systems", IRE Trans. IT-6, 386, (1960).

2. R.A.K. Said and D.C. Cooper, "Crosspath real-time optical correlator and ambiguity-function processor", Proc. IEE, 120, 423, (1973).

3. Robert J. Marks II, John F. Walkup and Thomas F. Krile "Ambiguity function display: an improved coherent processor", Appl. Opt. 16, 746 (1977).

4. David Casasent and B.B.K. Kijaya Kumar, "Optical image plane correlator for ambiguity surface computation", Appl Opt. 18, 1674 (1979).

5. Poohsan N. Tamura, Joseph J. Rebholz and T.C. Lee, "Ambiguity Map Generation with Coherent Optics", J. Opt. Soc. Am. 69 1451, (1979).

6. T.C. Lee, J. Rebholz, P.N. Tamura and J. Lindquist "Ambiguity processing by joint Fourier transform holography" Applied Optics, 19, 595, 1980.

7. P. Kellman, "Time Integrating Optical Processors", SPIE Proc. 185, (1979).

8. P.N. Tamura, J.J. Rebholz, O.T. Daehlin, and T.C. Lee, "Real Time Optical Computation of the Ambiguity Function", SPIE Proc. 241, (1980).

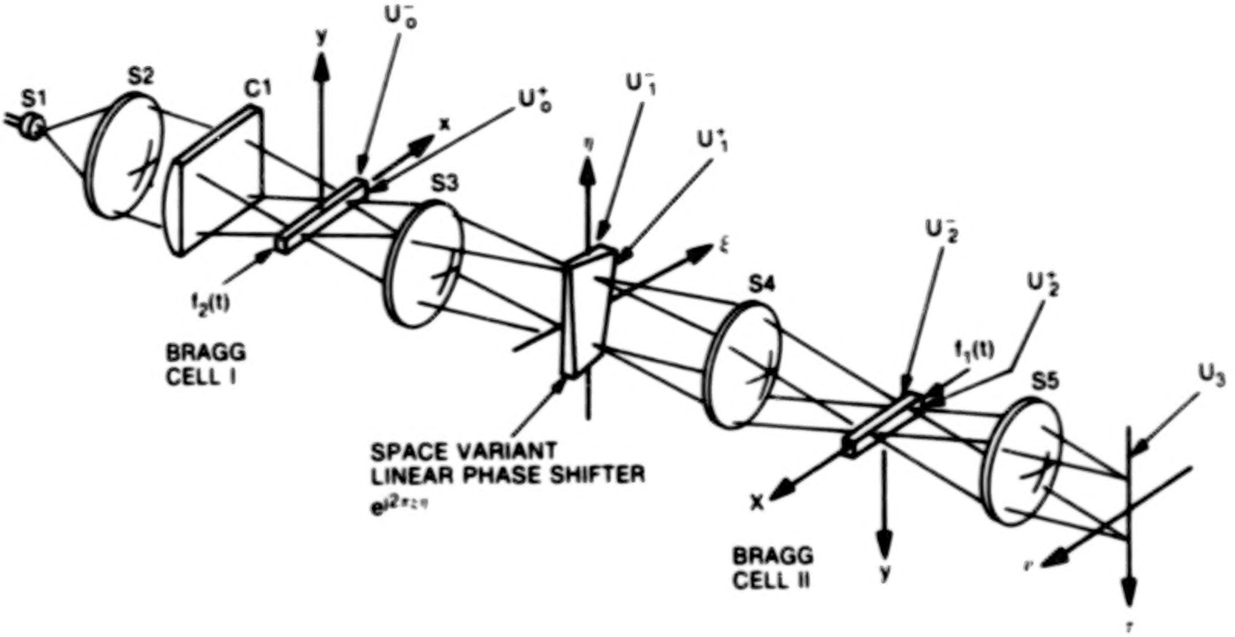

S1
S2
C1
y
U_0^-
x
U_0^+
S3
η
U_1^-
U_1^+
ξ
S4
U_2^-
U_2^+
$f_1(t)$
S5
U_3
$f_2(t)$
BRAGG CELL I
SPACE VARIANT LINEAR PHASE SHIFTER
$e^{i2\pi\xi\eta}$
x
BRAGG CELL II
y
ν
τ

Figure 1.- Optical layout for ambiguity function generation.

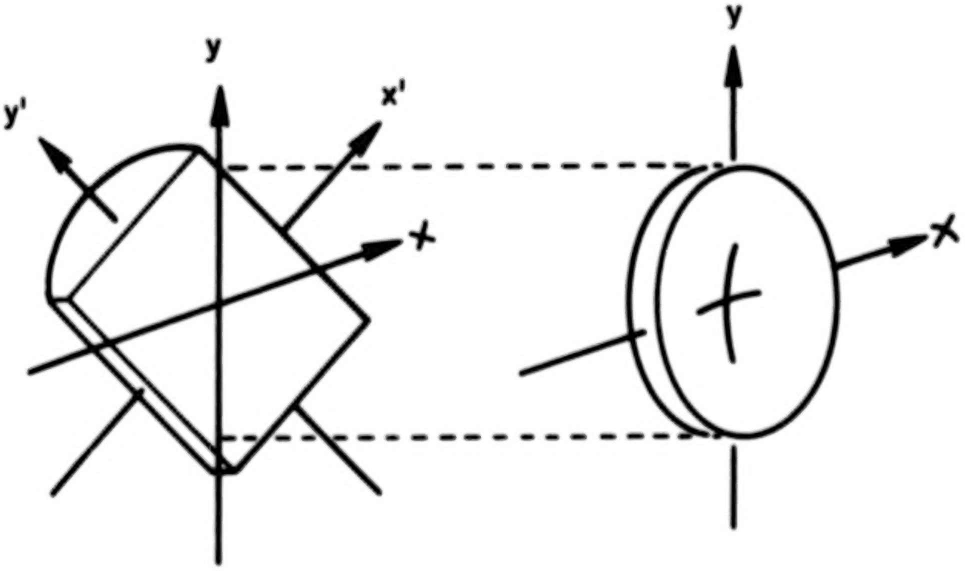


Figure 2.- Construction of the space variant linear phase shifter.

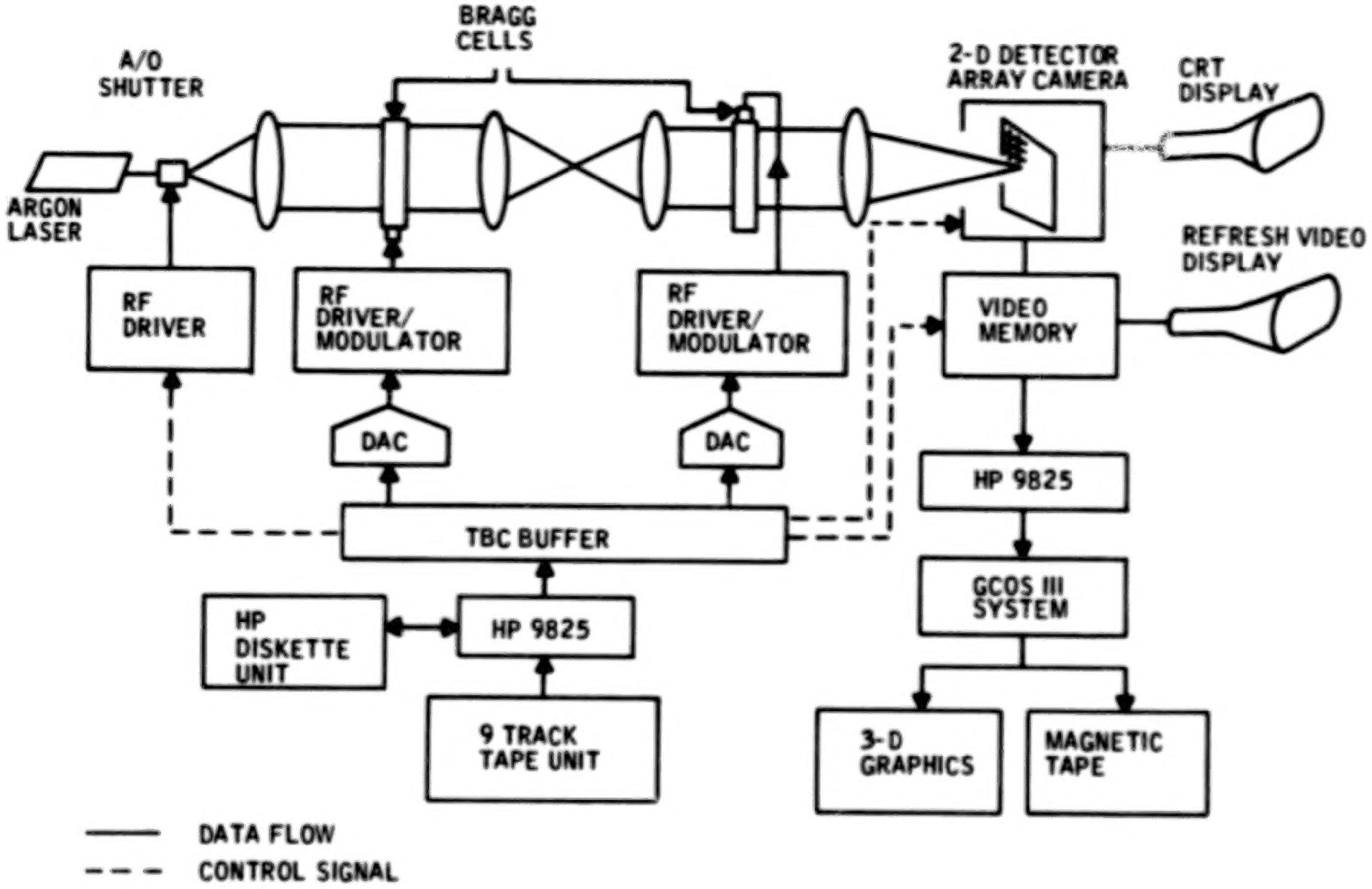

BRAGG CELLS
A/O SHUTTER
ARGON LASER
2-D DETECTOR ARRAY CAMERA
CRT DISPLAY
REFRESH VIDEO DISPLAY
RF DRIVER
RF DRIVER/ MODULATOR
RF DRIVER/ MODULATOR
VIDEO MEMORY
DAC
DAC
HP 9825
TBC BUFFER
GCOS III SYSTEM
HP DISKETTE UNIT
HP 9825
9 TRACK TAPE UNIT
3-D GRAPHICS
MAGNETIC TAPE
DATA FLOW
CONTROL SIGNAL

Figure 3.- Real time system block diagram.

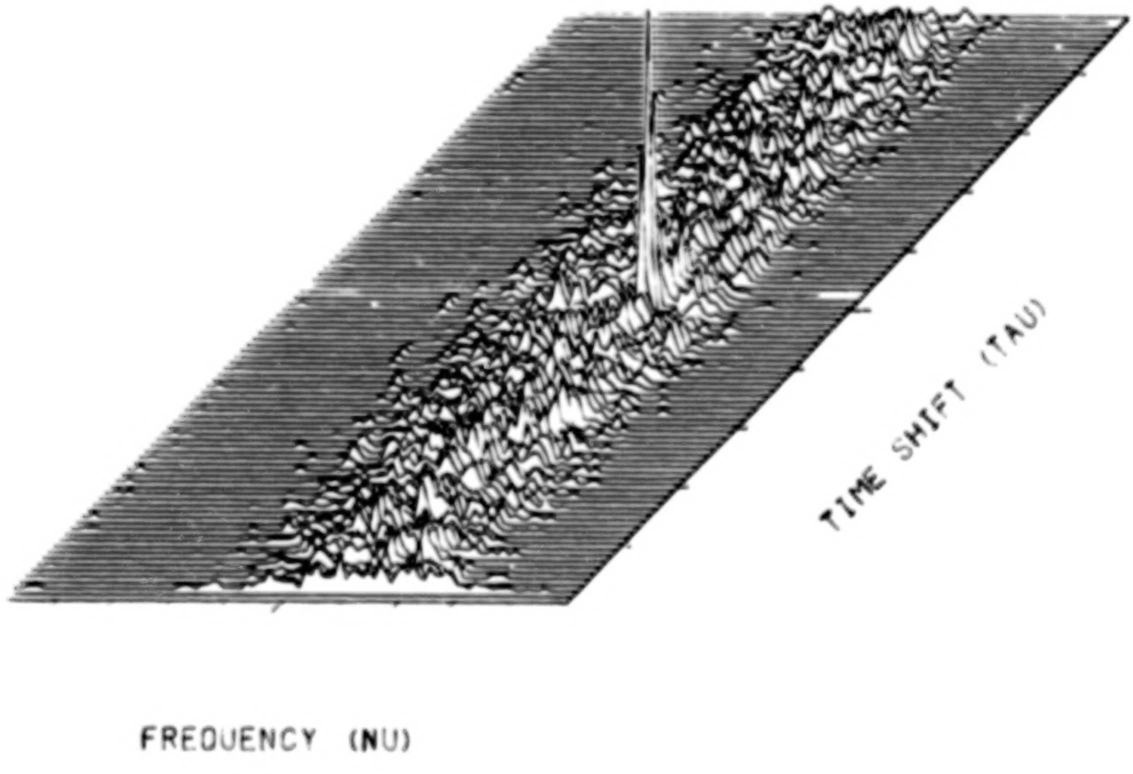


Figure 4.- Real time results, SNR = 0dB.

THE INTEGRATED OPTIC RF SPECTRUM ANALYZER

M.E. Pedinoff and T.R. Ranganath
Hughes Research Laboratories
3011 Malibu Canyon Road
Malibu, California 90265
(213) 456-6411

T.R. Joseph
Hughes Ground Systems Group
Fullerton, CA 92634

J.Y. Lee
Hughes Industrial Products Division
Carlsbad, CA 92008

ABSTRACT

The performance of a fully integrated rf optical spectrum analyzer (IOSA) is presented. This device exhibited a 3 dB bandwidth of 380 MHz with a diffraction efficiency of 5% (at 500 MW rf power). The fully integrated device operating at $\lambda_o = 0.82\mu$ exhibited the following performance figures: an acousto-optic bandwidth of 380 MHz, a 3 dB spot size of ∿8 MHz, a two-tone resolution of 8 MHz, and a linear dynamic range of >25 dB as observed visually on an oscilloscope. The resolution uniformity across the rf bandwidth is presently limited by the mechanical position errors of the semiconductor laser and the detector/CCD array in the focal planes of the collimating and focusing geodesic lenses.

All components used in the IOSA were first tested individually and reported here. A fully functional (with no dead detector cells) detector/CCD array exhibited a linear dynamic range of 30 dB, with a sensitivity uniformity of ±1 dB. A measurement of the undeflected zeroth order beam indicates geodesic lens insertion losses of $\lesssim$2 dB each. The acoustic transducer exhibited a 400 MHz 3 dB rf bandwidth.

INTRODUCTION

The use of guided-wave optics resulted in the development of a new class of one-dimensional optical processors that have a rigid, compact optical design and a low electrical drive power requirement. This type of processor is most easily applied to the analysis of rf signal spectra. In this paper we describe an integrated-optics circuit for real-time rf spectral analysis (ref. 1,2).

An integrated-optics optical spectrum analyzer consists of an injection laser diode, a thin-film optical waveguide, waveguide lenses, a surface acoustic wave (SAW) transducer, and a linear detector array. This device employs the interaction between a coherent optical wave and an acoustic wave driven by an input electrical signal to determine the power spectral density of the input. The basic design layout of an integrated-optic spectrum analyzer (IOSA) is shown in Figure 1. An incoming radar signal is mixed with a local oscillator so that the intermediate frequency (IF) is within the passband of the transducer. After amplification, the signal is applied to the SAW transducer. The resultant SAWs traversing the optical waveguide generate a periodic modulation of the refractive index of the waveguide mode. If a collimated, guided optical beam intersects the acoustic beam at the Bragg angle, a portion of the beam will be diffracted (deflected) at an angle closely proportional to the acoustic frequency, and with an intensity proportional to the power level of the input signal. The Bragg deflected light is then focused onto an array of focal plane detectors where each detector output becomes one frequency channel of the spectrum analyzer.

ACOUSTO-OPTIC INTERACTION

A diagram of the acousto-optic interaction region in a monolithic integrated optic structure is shown in Figure 2. The criterion for an efficient Bragg diffraction is that the interaction length L be sufficiently large so that the parameter Q is larger than 1; i.e.,

$$Q = \frac{L\lambda o}{n\Lambda^2 \cos \theta_i} > 1 \quad ,$$

where n is the waveguide modal index, Λ is the acoustic wavelength, λo is the optical wavelength, and θ_i is the incident angle of the optical beam defined in Figure 2. This criterion physically corresponds to an interaction length larger than the acoustic wavelength. The diffracted energy appears predominantly in one order.

When the acousto-optic interaction is phase matched in first order Bragg diffraction, the angle of incidence, θ_i, and the angle of diffraction, θ_d, are equal. The angular deflection δ (the angle between the incident and diffracted beam, or twice the Bragg angle) is given by

$$\delta = \theta_i + \theta_d = 2 \sin^{-1} \left(\frac{\lambda_o}{2n\Lambda}\right) \quad .$$

This angle is different for each acoustic wavelength or frequency. For a fixed incident optical beam direction, the direction of the acoustic wave must change with frequency. Within the small angle approximation, the deflection angle is linearly related to the acoustic frequency by

$$\frac{d\delta}{df_s} = \frac{\lambda_o}{nv_s} \quad ,$$

where f_s is the acoustic frequency, and v_s is the velocity of the surface acoustic wave. When the acousto-optic medium exhibits significant dispersion in v_s the deflection equation takes the form

$$\frac{d\delta}{df_s} = \frac{\lambda_o}{nv_s}\left(1 - \frac{f_s}{v_s}\frac{dv_s}{df_s}\right) \quad .$$

This indicates that nonlinear spacing of detector elements in the focal plane must be used for dispersive acoustic materials.

A complete Fourier transform analysis for this device was reported earlier (ref. 2). In this device the deflected optical beam generated in the Bragg cell modulator is proportional to the product of the incident optical beam wavefront, the accoustic wavefront, and the efficiency of the interaction. Therefore, as a result of the acousto-optic interaction, a fraction of the transmitted light beam acquires a new direction of propagation, corresponding to a bandwidth, df_s, centered at frequency f_s.

The output intensity distribution thus displays, at a position linearly proportional to the frequency of the modulating signal, a signal whose intensity is proportional to the intensity of the deflected beam (i.e., proportional to the spectral component of frequency f_s, multiplied by the efficiency of the Bragg modulator). For a modulator with a flat response over the frequency range of interest, the output is thus proportional to the intensity of the spectral component.

DESIGN PARAMETER STUDY

The optical circuit design of an IOSA device is constrained by the required rf frequency resolution and the optical resolution limit set by the allowed crosstalk between adjacent cells. The rf frequency resoltuion requirement sets the upper limit of the detector cell size S

$$S \le \frac{F\lambda_o}{nv_s}\Delta f \quad ;$$

while the optical resolution limit requires a minimum cell size determined by

$$S \ge \frac{g\lambda_o F}{nD} \quad ,$$

where D is the optical beam width, F is the output lens focal length, n is the effective guide index, and λ_o is the free-space wavelength. The resolution factor g is determined by the definition of optical focal spot resolution. These equations have been used to design an IOSA structure having 4 MHz rf resolution (ref.2) for an IOSA fabricated on an $LiNbO_3$ substrate, where the acoustic phase velocity is v_s = 3500 m/sec. For this device with a Gaussian beam truncated at the $1/e^2$ intensity points, the minimum beam width was 2.18 mm.

When the minimum optical beam width, D, is selected in the design, then the spot size of the optical beam at the detector array will just equal the detector cell size S. The adjacent channel crosstalk is defined as the total optical power collected by the cell adjacent to a cell corresponding to a given rf frequency divided by the total optical power collected by that rf frequency cell. This adjacent channel crosstalk ratio is plotted in Figure 3. The adjacent cell crosstalk is plotted as a function of DS. Thus, for an optical beam width of 2.18 mm, and a cell size of 8.0 μm (equal to the full lobe spot size at the focal plane), the crosstalk is -22.4 dB. If the lens quality is such that a lens diameter larger than D_{min} can be accommodated while diffraction-limited performance is still maintained, then the sidelobe intensity impinging on adjacent cells will be reduced. This will allow two closely spaced signals of significantly different energy levels to be detected and resolved. For a 4 MHz resolution and a cell size fixed by the full main lobe width of the minimum diameter beam (D_{min} = 2.18 mm), a 5.6 dB reduction in adjacent channel crosstalk results when the optical beam width is enlarged to three times this minimum value, or D = 6.54 mm. For this case, the adjacent channel crosstalk is -28 dB.

We have established that the value of $Sn/\lambda_o F$ necessary to achieve 4 MHz rf resolution is

$$\frac{Sn}{\lambda_o F} = 11.4 \times 10^{+2}\ m^{-1} \quad ,$$

where the ratio S/F for an $LiNbO_3$ system (n = 2.2) operating at λ_o = 0.82 μm is $S/F = 4.25 \times 10^{-4}$. Since the detector cell size, S, is set by detector array design considerations, the focal length of the Fourier transform lens is required to be $F = (S/4.25) \times 10^4$, which, for an 8-μm cell size, is 18.8 mm. If the optical beam diameter is chosen to be 7.4 mm, then the f/no. of the output transform lens is f/2.54. For this system, the adjacent cell crosstalk due to the optical beam diffraction pattern is below -28 dB.

PROPAGATION MEDIUM

An important issue in the development of an IOSA was the selection of the optimum substrate material system from the available candidates. The material selected for this device was a y-cut $LiNbO_3$ substrate with propagation direction along either the z axis or 21.8° from the z axis. $LiNbO_3$ is considered to be one of the best acousto-optic materials because of its low acoustic propagation loss, high electromechanical coupling constant, and the relative absence of velocity dispersion due to the optical guiding layer. With this substrate material the detectors are interfaced by direct butt coupling to the $LiNbO_3$ waveguide. For y-cut crystals, $LiNbO_3$ substrates have different SAW characteristics for propagation along the z axis, and for propagation 21.8° off the z axis (ref. 3). An analysis of acousto-optic modulation efficiency indicated that there is no significant difference between these orientations. From the optical standpoint, the z-oriented acoustical transducers minimize the substrate anisotropy and are preferred.

The waveguide on the polished y-cut $LiNbO_3$ substrate is formed by in-diffusion of electron-beam-evaporated Ti (ref. 4). A fabrication sequence is used in which the important parameters (such as surface quality, optical flatness, cleanliness, uniformity, and adherence of the metal films to the substrate) are controlled throughout the process. Single-mode optical waveguides are formed by diffusing 200-Å-thick Ti at 950°C for 4 hours. The diffusion depth of the resulting waveguide is about 2 µm.

The propagation loss in $Ti:LiNbO_3$ waveguides at the He-Ne laser wavelength (λ = 6328 Å) is typically less than 1 dB/cm. The waveguide loss for a TE polarization wave (He-Ne laser wavelength) was experimentally determined to be 0.85 dB/cm (ref. 2).

In-plane scattering of the optical waveguide is an important parameter in determining the dynamic range of the IOSA. He-Ne laser light was used to excite the waveguide mode and the strength of the in-plane scattering was measured by scanning the m-line extracted by a coupling prism. The scattering intensity shown in Figure 4 is -40 dB for a scattering angle of 0.6° in the waveguide, corresponding to an acousto-optic deflection angle due to an acoustic frequency of 115 MHz. In-plane scattering levels of -49 dB and 57 dB of the peak intensity were measured at an in-plane scattering angle of 1.3° (corresponding to deflection by a SAW having a frequency of 270 MHz) on each side of the main beam. When the IOSA is operated at the GaAlAs laser wavelength, the in-plane scattering level is further reduced.

LASER CHARACTERIZATION

To meet the frequency resolution specification, the total output spectral width of the laser should be much less than 40 Å under operating conditions. Ideally, the output should be in a single transverse and single longitudinal mode with spectral width <1 Å. The lasing aperture should be as small as possible to simulate a point source so that the problem of collimating the laser output in the thin-film waveguide can be simplified. Based on these considerations, we selected the Hitachi GaAlAs buried-heterostructure injection laser diode as our source. The lasing aperture of the device is ~1 µm in diameter with nearly isotropic divergence. In this program we tested two special-order lasers that have an emitting edge flush with the heat sink edge. They are double-heterojunction lasers, Model HLP2400, Serial Nos. 3831 and 3834. The far-field intensity distribution in the two principal orthogonal directions (perpendicular and parallel to the junction) was measured using a photomultiplier and a slit with an effective angular resolution of 0.115°. Figures 5 and 6 show the far-field intensity profile measured at HRL. The beam divergence in the plane parallel to the junction is less than 20° for both lasers. In the plane normal to the junction, the beam divergence is 30° (Serial No. 3834) and 37° (Serial No. 3831), respectively.

There is an upper limit in the laser power coupled into the waveguide at which the detector cell is saturated. The laser power requirement can be estimated based on the following values: a saturation charge density of the detector cells of 6.0×10^6 electrons/cell, a detector quantum efficiency of 0.25, a maximum acousto-optic deflection efficiency of 0.05, an optical propagation loss of 9 dB, and a detector integration time of 3.0 µsec. The calculation shows that about 600 µW of laser power has to be coupled into the waveguide.

WAVEGUIDE GEODESIC LENSES

Waveguide geodesic lenses are waveguide-covered depressions (or protrusions). The curvature of the depression generates a difference in geometric path length for rays traversing different portions of the structure; this produces a focusing effect. The optical path length at the center of the depression is longer than at the sides; and the phase fronts are curved toward the lens axis beyond the depression. If the depression is of a particular aspheric shape, or if the wavefront emerging from the line is appropriately compensated by external overlayers, the focusing will be aberration-free (except for field curvature).

The design of the aspherical depression profile for a diffraction-limited geodesic lens is based on the equivalent-lens principle (ref. 5). From a generalized Luneburg lens profile calculated by numerical computation, one can derive the equivalent aspherical geodesic lens profile. The geodesic lens has no chromatic aberrations because of its uniform refractive index. This is independent of dispersion in the material refractive index. The focal length of the geodesic lens is independent of the modal index, and it can be used in both single- and multi-mode optical systems.

Diffraction-limited geodesic lenses have been demonstrated in two different waveguide structures: polyurethane film-glass substrate (ref. 6) and Ti in-diffused $LiNbO_3$ waveguide (ref. 7). The focal-length measurements (ref. 7) indicate that the lens is aberration free up to about 90% of the aperture. Two waveguide geodesic lenses are used in the IOSA: one collimation lens and one Fourier transformation lens. Both lenses are identical, and their parameters are listed in Table 1. Since the aspherical lens profile has a smooth transition region that takes up 15% of the lens aperture, the useful lens aperture is 7.4 mm.

FABRICATION OF ASPHERICAL DEPRESSIONS

The fabrication of geodesic lenses involves two stages: grinding and polishing the lens depression with high precision. Ultrasonic impact grinding has been selected for both of these operations. The key problem appearing in the optical polishing is to obtain an optical surface without destroying the complex lens profile, in particular, the edge rounding region.

Ultrasonic impact grinding employs a tool bearing the shape of the designed depression. Ultrasonic energy transmitted through the tool agitates the abrasive slurry, and material is removed from the sample surface. The ultrasonic impact grinding technique is capable of replicating the tool shape within a 1/4-μm tolerance (ref. 6). The operation requires more than one tool to reduce the effects of tool wear.

Ultrasonic impact grinding is a complicated process governed by many parameters. We have empirically determined the optimum values of tool shapes, tool materials, grinding pressures, stroke amplitudes, workpiece materials, slurry speeds, abrasive concentrations, and grit sizes for our specific lens designs. Double lens pairs have been ground on 6.35 cm x 3.81 cm x 0.318 cm (x-z-y) $LiNbO_3$ substrates. After polishing and diffusion of Ti to form planar waveguides, we experimentally determine the location of the focal planes of the lenses and polish the edges to coincide with the focal planes.

LENS TOLERANCE ANALYSIS

Recently, Betts et al (ref. 8) published expressions for tracing rays in rotationally symmetric geodesic lenses. These expressions have been used to calculate the effects of tool wear, and lens depth.

The profile used in the calculations is for an f/1.5 lens. Some of the results of these calculations are given in Figures 7 and 8. Figure 7 shows the focal length increase, with percentage depth for both the tool wear analysis and the decrease in depth while maintaining the actual profile (surface removal).

For an f/1.5 lens, the diffraction-limited spot size is .58 µm at a wavelength of 0.8 µm in $LiNbO_3$. For a 1-cm-diam lens, this corresponds to 3×10^{-4} units in Figure 8. Hence, to maintain diffraction-limited performance, a depth change of ∿1% can be tolerated if 90% of the original useful aperture is used. For greater tool wear, a smaller aperture is required to maintain a diffraction-limited performance. This was experimentally demonstrated with a lens fabricated in $LiNbO_3$, where an 80 % aperture gave near-diffraction-limited performance.

From the experimental and calculated results, we concluded that in any system fabricated using geodesic lenses the precise location of the focal plane could only be determined after the lens fabrication. The focal position is very sensitive to fabrication tolerance, but the focal spot size remains diffraction limited within achievable fabrication tolerances.

SAW TRANSDUCER DEVELOPMENT

During this phase of the program we developed a 500-MHz bandwidth, 1-GHz operating frequency SAW transducer suitable for use in the IOSA, and we acoustically characterized the $Ti:LiNbO_3$ structure. Experiments were carried out to measure propagation loss, velocity dispersion, and acousto-optic deflection efficiency on various $Ti:LiNbO_3$ substrates. A modified chirp transducer (MCT) was able to achieve acousto-optic deflection bandwidths exceeding 400 MHz for both HeNe and GaAlAs laser wavelengths. This is the largest acousto-optic deflection bandwidth reported to date for a single transducer at either laser wavelength.

The MCT is similar to a conventional chirp transducer in that it has electrode spacings that vary smoothly from one end of the transducer to the other. This ensures that a portion of the transducer electrodes is synchronous at all frequencies within the passband. The chirp transducer response is relatively smooth over its bandwidth.

The total number of electrodes in the chirp transducer can be changed to facilitate impedance matching with virtually no effect on the shape of the output spectrum. In fact, the number of electrodes can be set so as to yield the ideal transducer loss with no external matching circuit whatsoever. In the case of the chirp transducer, each interelectrode gap has a characteristic ("instantaneous") synchronous frequency and it can be tilted to satisfy the Bragg condition at that frequency. Figure 9 shows the modified chirp transducer configuration with tilt angle varying smoothly over the transducer band. At any frequency within the band, a small number of electrodes are in quasi-synchronism, while contributions from the remaining electrodes add in virtually random phase. For this reason, SAW generation is confined to a single "active" region that slides along the transducer as the drive frequency is changed. The number of excitation electrodes N_a comprising an "active" section for a linear FM chirp transducer is given by

$$N_a(f) = \frac{f}{f_o} \; \frac{N_{TOT}}{\sqrt{\Delta f \Delta \tau}} \quad ,$$

where $N_{TOT} \approx 2f_o \, \Delta\tau$ is the total number of electrodes, f_o is the midband frequency, and $\Delta\tau$ is the transducer length divided by the SAW velocity. The total number of electrodes can be varied to satisfy convenient impedance-matching conditions. When the transducer aperture is fixed by acousto-optic interaction length considerations, N_{TOT} can be increased until the midband capacitive reactance reaches 50 Ω. At this point, the transducer can be coupled to a 50 Ω generator without any tuning. In addition, losses due to parasitic resistance in the metal film can be reduced by making N_a large, since the electrode resistance term is inversely proportional to the number of active electrodes.

The transducer pattern ultimately used in the development of the IOSA was a modified chirp transducer operating at 600 MHz with a 400-MHz bandwidth. This transducer, shown schematically in Figure 9, consists of 180 electrodes with $N_a = 21$ arranged in a two-track "dog-leg" structure to improve the impedance matching characteristics. The total aperture is 76 wavelengths, and the total tilt angle variation is 1:7°, since this design covers the frequency range from 350 to 900 MHz. The transducer photomask pattern with non-parallel electrodes was programmed on a computer-controlled optical plotter and photographically reduced in 3 stages. The devices were then built using contact optical photolithography. In all cases, the transducers were etched from 500-Å-thick aluminum metalizations.

PERFORMANCE

Data was obtained on acoustic conversion efficiency, propagation loss, velocity dispersion, and acousto-optic deflection efficiency.

The 600-MHz MCT pattern was initially fabricated on untreated yz $LiNbO_3$. Measurements were made of the insertion loss of this transducer in a non-dispersive delay-line configuration and are shown in Figure 10. The measured conversion efficiency agreed quite well with the theoretical value of 13 dB. The insertion loss varies from 14 dB at the lower edge of the band to 12 dB in the middle. Varying the length of the gold wire leads causes the passband to tilt by as much as 5 dB. Excellent agreement between the theoretical and the measured response was achieved.

Using test delay lines fabricated on Ti-diffused $LiNbO_3$ samples, measurements were made of acoustic propagation losses. The resulting data is shown in Table 2. To obtain this data, it was assumed that the measured insertion loss of the untreated samples consisted of the sum of the conversion efficiency and the theoretical propagation loss predicted by Szabo and Slobodnik (ref. 9).

There is great disparity in the propagation loss data obtained. An attempt was made to correlate the attenuation values with the temperature profiles used during the Ti diffusion. The times at each temperature are also shown in the table.

Sample N-92, which exhibited essentially theoretical propagation loss values, was heat treated at 650°C for 2 hours and at 950°C for 5 hours. The higher losses measured for sample N-95 indicate that the time at both 650°C and at 950° C are critical in determining the propagation loss of the samples.

The layered Ti:LiNbO$_3$ structure is expected to exhibit some acoustic velocity dispersion. It is important to characterize the amount of dispersion because the frequency dependence of the optical deflection angle δ and hence of the detector spacing is dependent on velocity.

The dispersion of the acoustic time delay was measured for crystals with differing diffusion temperature profiles. The results are given in Table 3. These results were obtained from a computer-controlled rf network analyzer that measured the insertion phase of a delay line consisting of two 600-MHz MCTs and performed a second-order least-squares fit. The values of 1.6×10^{-5}/MHz measured for these devices corresponds to a change in time delay of 0.64% for a 400-MHz bandwidth. However, of more importance is the error in position of the detector cells because they are uniformly spaced. The fractional position error of the high-frequency end of the detector can be calculated relative to the low-frequency end. The result is that for a 400-MHz bandwidth centered at 600 MHz, the error is 1%, or 1 detector cell width.

Acousto-optic deflection data was obtained for devices consisting of Ti:LiNbO$_3$ substrates with 600-MHz MCT transducers. The setup shown schematically in Figure 11 was found to be the most convenient and the most accurate method for obtaining this data. The light beam from the laser is coupled into the waveguide using a rutile prism coupler. The guided optical beam traverses the acoustic path, and the undeflected and deflected beams emerge from the waveguide at the edge of the crystal. There they are imaged on solid-state detectors using a cylindrical lens. The deflection efficiency is then determined by taking the ratio of the two optical intensities. The SAW transducer is driven by an rf oscillator and amplifier. The incident rf power can be monitored using a power meter. The primary advantage of this technique is that no output prism coupler is required. The main disadvantage of this technique is that a chip-free polished edge must be obtained on the output edge of the LiNbO$_3$ substrate. Otherwise, the light in the waveguide is scattered and cannot be properly imaged on the detectors.

Figures 12 and 13 show acousto-optic deflection efficiency versus frequency measured with 100 mW of rf power incident on the SAW transducer.

Figure 12 shows the response obtained using a GaAlAs double-heterojunction laser with a TE polarization, and with a wavelength of 0.83 µm. The 0 dB here corresponds to a 1.6% deflection efficiency; the transducer bandwidth is 420 MHz.

As proof of the effectiveness of the tilted fingers to steer the acoustic beam and affect the bandwidth, a measurement was made on the deflection coefficient when the laser was incident from the opposite side. This is equivalent to changing the sign of the direction in which the electrodes tilt. The result is a response that is significantly narrowband compared with that shown in Figure 12. Using Equation (15) (of Ref. 10), the shape of this response can be calculated; it is shown in Figure 13, along with the curve measured. The two curves agree quite well and the response is significantly narrowband relative to the parallel electrode case. The slight displacement in frequency between the two curves results from the experimental uncertainty concerning the frequency at which the Bragg condition was exactly satisfied.

DETECTOR ARRAY AND CCD READOUT

The detector array of the IOSA is designed to detect SAW deflected laser light, to provide information on incoming rf frequency, and to read out with a charge-coupled-device shift register.

The CRC 139 detector array is made of a single row of 100 silicon MOS photodetectors and 4 output CCD shift registers located on both sides of the detectors. These small photodetectors are made of MOS capacitors with semi-transparent polysilicon gates. The CCD shift register is fabricated using polysilicon gate buried-channel CCD technology to reduce noise and increase the rate of charge transfer. The CCD is fabricated with n-channel CCD technology to take advantage of higher electron mobility. The schematic cross section of a buried p-channel MOS photodetector compatible with the above n-channel CCD technology is shown in Figure 14.

When a bias voltage is applied on the gate of the MOS capacitor, a depletion region is formed under the polysilicon gate. Electron-hole pairs generated by photons penetrating the silicon substrate will be separated by the electric field in the depletion region. The electrons collected at the Si-SiO$_2$ interface become the signal charge.

The primary performance goals of the detector array were:

- An rf frequency resolution of 4 MHz
- A dynamic range of 40 dB
- A total sampling time of <3 μsec.

The detector cell width S is constrained by the required rf frequency resolution and by the allowed optical crosstalk between adjacent cells discussed earlier. Using the parameters associated with Figure 3, the detector cell width defined by frequency resolution requirements is $S \leq 4.26 \times 10^{-4}F$, and the cell width defined by optical resolution requirement is $S \geq 1.32 \times 10^{-4}F$. A small focal length F corresponds to a small detector size. The detector width, however, is restricted in the lower limit by detector-CCD fabrication and coupling considerations, and by dynamic range requirements. The detector coupled to CCD shift registers is shown in Figure 15. To facilitate the charge transfer, it is desirable to have equal width for both the detector and the CCD cell. It is difficult to make the width smaller than 16 μm per bit or 8 μm per detector for a 2 detector to 1 bit dual multiplexer. There are channel stops between the detector cells to isolate the cells from one other. The minimum channel stop width that can be made on a mask is 1.5 μm and will be 3.5 μm wide after processing. With a spot size of about 2.5 μm, the detector size should be at least 7 μm to contain the full light spot. To satisfy the dynamic range requirement of 40 dB requires a rather large detector storage area. Any further reduction in detector width will make the detector length even longer and slow down charge transfer. Based on these considerations, we chose a detector width of 8 μm for the Fourier transform lens with a focal length of 1.88 cm. This gives the required frequency resolution of 4 MHz.

A dynamic range of 40 dB is required for system applications. The dynamic range is defined as the range between detector saturation due to strong signal and the overall noise level. The determination of expected noise level is, therefore, very important for designing the full charge capacity of both the detector and CCD unit cells.

The usual noise sources of a detector/CCD array can be separated into three categories: those associated with the input, the integration and transfer, and with the output. These noise effects are listed in Table 4 with special emphasis given to the comparison of noise characteristics of surface- and buried-channel CCDs.

The conclusion that the noise level of a surface-channel CCD is about four times higher than that of a buried-channel CCD has led to the selection of the buried channel device for this program. To achieve a dynamic range of 40 dB, or 10^4 to 1, the charge capacity of both the detector and the CCD storage well must be at least 10^4 times greater than the noise level. The noise level of a buried-channel CCD and its output amplifier is about 210 electrons. A full charge capacity of 6×10^6 electrons is thus designed for this array. Assuming a charge density of 1×10^{12} electrons/cm^2, this results in a detector of

$$\frac{6 \times 10^{6} \text{ electrons}}{1 \times 10^{12} \text{ electrons/cm}^2} = 6 \times 10^{-6} \text{ cm}^2 .$$

With an effective detector width of 4.5 µm, this requires a detector length of at least

$$\frac{(6 \times 10^{-6} \text{ cm}^2)}{(4.5 \times 10^{-4} \text{ cm})} = 1.33 \times 10^{-2} \text{ cm} = 133 \text{ µm} .$$

Therefore, the detector size is designed to be 8 µm by 150 µm.

The 100 detectors should be sampled as fast as possible, preferably within 3 µsec. Therefore, the detector, the CCD shift register, and the output amplifier are all required to have high-speed operation. Three basic mechanisms are responsible for carrier transfer in a CCD: thermal diffusion, self-induced drift, and fringing fields. A summary showing thermal diffusion time constants, fringing field, and carrier mobility is shown in Table 5. An epi concentration of $5 \times 10^{14}/\text{cm}^3$ (30 Ω-cm) and a potential difference of 10 V between adjacent electrodes are assumed. Also listed in Table 5 are t_4, the time to achieve a transfer efficiency of four 9's or 0.9999, and f, the maximum operating frequency assuming a 2-phase clocking.

With a readout scheme as shown in Figure 15, each of the CCD registers will have a maximum of 100 transfers between detector and output amplifier. If a 3-µsec sampling time is desired, a transfer time t_4 of less than 3×10^{-8} sec would be required. Table 5 shows that CCDs with an electrode length of 7.5 µm or less have a very good chance of meeting this goal. The electrode length chosen in the present design is 6 µm.

IOSA DEVICE TESTING AND PERFORMANCE

The IOSA is a complex device and extensive testing is needed for performance evaluation. Looking at the close-up of a fully assembled IOSA in Figure 16 we can clearly discern the detector/CCD chip at the top end, the two geodesic lenses in the middle, and the semiconductor laser at the lower end of the $LiNbO_3$ integrated optic chip. Two sets of SAW transducers can be seen in the middle of the $LiNbO_3$ chip on either side of the axis of the crystal. The significant features of the IOSA's performance are:

- Acousto-optic bandwidth
- Single-tone rf resolution
- Single-tone dynamic range
- Two-tone rf resolution
- rf pulse measurements.

Acousto-Optic Bandwidth

The IOSA device depends on acousto-optic Bragg deflection for its successful operation. The bandwidth associated with the acousto-optic interaction depends on the rf bandwidth of the transducer for efficient SAW generation, as well as k-vector matching of the optical and surface acoustic waves. But a purely rf characterization of the transducer reveals only a part of the picture. In order to evaluate the acousto-optic response we make use of an experimental arrangement, illustrated by the schematic in Figure 17. To obtain the optimum acousto-optic response we need to arrive at the correct position for the semiconductor laser. This is decided by measuring the acousto-optic response corresponding to different transverse positions of the semiconductor laser with respect to the axis of the collimating lens. In this measurement a 1 kHz amplitude modulated rf signal is amplified and applied to the SAW transducer of the IOSA. The directional coupler between the rf amplifier and the SAW transducer facilitates measurement of the incident and reflected rf power levels. The deflected optical spot at the output edge of the crystal is imaged onto a photomultiplier. The beam splitter arrangement allows monitoring of the output edge of the crystal on a video monitor. The signal from the photomultiplier, together with the 1 kHz modulating signal from the wavetek oscillator, is fed to the signal and reference channels of a PAR 124A lock-in amplifier. The analog output from the lock-in is used to drive an XY-plotter. Because of the lock-in measurement technique employed, any scattered light from the substrate does not cause problems. With this experimental arrangement we have been able to rapidly measure the acousto-optic bandwidth for each position of the semiconductor laser.

The optimum acousto-optic response, measured with an HLP 1400 laser coupled to the IOSA chip, is shown in Figure 18. The 3 dB bandwidth is ∿380 MHz. The acousto-optic deflection efficiency corresponding to the 0 dB level at a frequency of 615 MHz and 100 mW, was measured as the ratio of the deflected spot signal to the undeflected spot signal. The deflected spot signal was arrived at by substrating out the signal due to any scattered background. For the two samples tested we have observed efficiencies of 0.83% (sample 4 with HLP 3400 U laser), and 0.97% (sample 3 with HLP 1400 CSP laser), respectively. The irregular structure observed in the acousto-optic response of Figure 18 is due to rf mismatch which arises from the coaxial cable (RG57, BNC) between the output of the directional coupler and the SAW transducer. This has been checked by monitoring the reflected power, keeping the incident power at a fixed value of P_{avg} ∿100 mW, and measuring the acousto-optic response over the frequency range 400 MHz to 900 MHz. In Figure 19 we have plotted the acousto-optic response, as well as the net rf power flowing into the SAW transducer/coaxial cable pair versus frequency. The correspondence between the peaks and valleys of the acousto-optic deflection efficiency with the peaks and valleys of the net power into the SAW transducer confirms this conclusion.

Single-Tone RF Resolution

The IOSA device was designed to have an acousto-optic 3 dB bandwidth of 400 MHz, with a center frequency of 600 MHz and a single-tone resolution of 4 MHz. Acousto-optic response measurements discussed earlier indicate that the Bragg cell design does indeed yield a 3 dB response of $\sim$380 MHz, very close to the $\geq$ 400 MHz rf bandwidth. The single-tone resolution of the device depends on a number of factors such as the gap between the laser emission facet and the input end of the $LiNbO_3$ crystal, the positions of the focal planes of the collimating and Fourier transforming lenses with respect to the edges of the $LiNbO_3$ crystal, SAW transducer alignment, and the distance between the output end of the $LiNbO_3$ crystal and the surface of the detector/CCD chip.

Even though the acousto-optic response was less than optimum, single-tone frequency measurements performed on the device are nevertheless interesting.

A single-tone measurement carried with a 725 MHz 100 mW rf signal is shown in Figure 20. The detector/CCD chip is organized in a way that data is read out in two channels as strings of 50 pixels each. In order to determine the frequency resolution we need to have data from both the output channels properly interleaved on the display. The photograph in Figure 20 shows output of all the 100 pixels in proper time sequence. The upper photograph is an oscilloscope trace of the 100 detector/CCD array. Referring to Figure 20, the straight horizontal part of the trace on the left-hand side defines the zero signal level. The pedestal seen in all of the other pixels in Figure 20 is due to in-plane scattered light. The pixel corresponding to the rf signal of 725 MHz has been highlighted. The adjacent cells that are 4 MHz away by design exhibit a signal half as strong as the pixel corresponding to the peak. The 3 dB frequency resolution is therefore 8 MHz. The lower half of Figure 20 illustrates the display on a regular electronic spectrum analyzer when the same 725 MHz rf signal is incident on it. As a second example, the photograph in the upper part of Figure 21 illustrates the output from one half of the detector/CCD array when a 705 MHz, 100 mW rf signal is incident on the SAW transducer of the IOSA. Once again, the lower part is the display of a conventional spectrum analyzer with the same 705 MHz, 100 mW rf signal incident on it.

Single-Tone Dynamic Range

The dynamic range of response is an important parameter in the evaluation of the IOSA. As outlined earlier, in-plane scattered light appearing at the output of the $LiNbO_3$ crystal results in a fixed pattern of noise impinging on the detector/CCD array. The resulting pedestal of background light is clearly illustrated in Figure 20. Dynamic range measurements were carried out with respect to the fixed background pattern. Measurements were carried out without resorting to a lock-in type of phase sensitive detection. With the output of the detector/CCD chip displayed on an oscilloscope, we read signals off the screen corresponding to each value of the rf drive signal. Also, to get an estimate of the cross-talk figure for our device, we measured the signal appearing in a pixel which was next to the nearest neighbor, corresponding to a frequency gap of 8 MHz. The result of one such measurement is shown in Figure 22. The detector/CCD clocking was carried out at 333 kHz, with the rf at 729 MHz and an $I_{laser} = 21$ mA. The signal in the main pixel saturates for rf power greater than about 200 mW, resulting in a dynamic range of 20 to 23 dB, with the minimum visually discernable signal level at about 1 mW (rf). The receiver sensitivity for the IOSA as determined by observing the detector/CCD output on an oscilloscope would be 1 mW (of rf power). The signal in the next to nearest neighbor (8 MHz array) appears to be a factor of 4 smaller than the main pixel. This measurement is again in agreement with our 3 dB resolution figure of 8 MHz (Figure 20).

The dynamic range is currently limited by saturation of the detector at the high signal end, and by the presence of the fixed pattern background signal at the lower end. Steps such as improved crystal handling, improved waveguide fabrication, and polishing of the crystal surface after waveguide formation can be pursued to reduce in-plane scattering of light.

Two-Tone RF Resolution

Functionally, the IOSA is expected to receive and analyze multiple rf signals for their frequency and amplitude characteristics. An important parameter in analyzing more than one rf component at a time is the minimum spacing between adjacent signals below which the device fails to resolve. The two-tone resolution measurement attempts to quantify the device performance in regard to simultaneous multiple signal spectral analysis. In this device design, the spacing between adjacent detector pixels corresponds to an rf increment of 4 MHz. If the achieved rf resolution were such that a significant response was observed in detector pixels adjacent to the intended pixel, two-tone resolution would suffer. Single-tone resolution measurements discussed earlier (see Figure 20) have shown that the strength of the signal in an adjacent detector pixel 4 MHz away is only a factor of 2 smaller. Assuming a simple Gaussian spot shape impinging on a detector element, we would not expect the device to resolve a pair of adjacent rf signals of equal strength and 8 MHz separation. Yet our measurements with the device indicate that a pair of equal strength rf signals 8 MHz apart are resolved. Oscilloscope displays of the output of the detector/CCD chip, when the SAW transducer is driven by two rf signals of 100 mW intensity separated by 8 MHz, is illustrated in Figure 23. In Figure 23(a) the output of the 100 element detector/CCD chip is shown in proper time sequence. In Figure 23(b) the outputs from the two shift registers are displayed separately. The rf spacing between adjacent points in Figure 23(b) is 8 MHz, while the spacing between adjacent points in Figure 23(a) is 4 MHz. In Figure 23(c) and (d) we show time expanded details of the displays in parts (a) and (b). The rf spectrum of the signal used to generate the output in Figure 23(a) is shown in Figure 23(e). The central signal at 689 MHz is 20 dB smaller than its two neighbors at 685 MHz and 693 MHz, respectively. Hence, the full 200 mW of rf power is distributed equally between the two signals at 685 MHz and 693 MHz. An examination of the display in Figure 23(a) shows that the pixel corresponding to 689 MHz exhibits a response 0.7 dB below the pixels at 685 MHz and 693 MHz. Thus, two equal strength signals separated by 8 MHz are clearly resolved.

In order to characterize the two-tone resolution of the device, we carried out a number of other measurements. Similar displays were obtained when the spacings between the two 100 mW rf signals were 12 MHz, 16 MHz, 24 MHz, and 40 MHz, respectively. Resolution of the two rf signals in each case was very clear, as shown in Figure 24, with a frequency offset of 12 MHz.

An interesting feature of the two-tone measurement is concerned with the measurement of the dynamic range of a signal in the presence of a strong signal nearby. Referring to Figure 20, we find it necessary to move 20 MHz (5 or 6 adjacent detector pixels) to achieve a cross figure of -20 dB. This means that in the presence of a strong signal, a weak signal would not be clearly resolved (with a cross-talk figure of -20 dB) unless it were ≥ 20 MHz away. The two-tone dynamic range is expected (if $\Delta f \geq 20$ MHz) to be better than 10 dB for the weak signal. In order to improve the two-tone dynamic range we would need to improve the frequency resolution.

rf Pulse Measurements

The IOSA device is expected to receive and analyze rf pulses of short time duration. If a laser with a divergence angle of 11.45° is used with this optical system (with a focal length of 20 mm), it generates a collimated optical beam 4 mm wide. Combining the 4 mm optical beam width with a surface wave velocity of 3.5×10^5 cm/sec yields an acoustic aperture of 1.15 µsec. At any given instant in time the acousto-optic interaction can only analyze an acoustic wave train that is 1.15 µsec long. Because of the integrating nature of the detector/CCD chip, all light impinging on a given detector pixel during one integration period contributes to the signal packet. An acoustic aperture of 1.15 µsec corresponds to a frequency resolution of 0.85 MHz, and there does not appear to be an advantage in integrating signals for periods longer than the acoustic aperture. For analyzing pulses, it would be advantageous to carry out the signal integration process for the duration of one acoustic aperture. As the integration time is reduced, one must increase the power output of the semiconductor laser by a proportionate amount in order to maintain the large signal handling capacity of the device. At the same time, laser power scaling has a definite upper limit due to the limited power output (10 to 15 mW) of single mode semiconductor lasers.

With the present device we have operated the detector/CCD chip at a clocking frequency of 333 kHz, with an integration time of 192 µsec. With a laser power output of 1.5 mW and a total insertion loss of 17 dB from the laser facet to the output of the crystal, signal saturation appeared at rf drive power levels of 200 mW. We have carried out a set of calculations to arrive at different sets of operation alternatives for our IOSA. The results are summarized in Table 6. The last row in Table 6 indicates that with a 15 mW laser output and an integration time of 0.96 µsec, signal saturation should occur at an rf power level of 1000 mW. These are all numbers that appear reachable at this time.

To highlight some of the above issues, we present rf pulse measurement data in Figure 25 for the case when the integration time was 38 µsec (CCD clocking frequency of 1.67 MHz) and pulses of various duration were applied to the SAW transducer successively. The pulse widths were 38 µsec, 19 µsec, 15.2 µsec, 11.4 µsec, 7.6 µsec, and 3.8 µsec. The rf (710 MHz) power level was set at 100 mW under cw conditions, and the modulating signal's pulsewidth was varied to achieve the different duty factors. The signal level is reduced successively in direct proportion to the duty factor, as shown in the displays in Figure 25(a) through (f). For purposes of establishing the background level, we show the detector/CCD output in Figure 25(g) with no rf power applied to the transducer. Finally, in Figure 25(h) we have included the spectrum of the modulated signal (duty factor 0.2) as seen on a regular electronic spectrum analyzer. The displays in Figure 25(a) through (g) highlight the problems involved in analyzing rf pulses that are small fractions of the detector/CCD integration times.

CONCLUDING REMARKS

The results of measurements made on an integrated optic rf spectrum analyzer are reported. The performance of the device acousto-optic bandwidth, single-tone rf resolution, two-tone rf resolution, single-tone dynamic range, two-tone dynamic range, and single-tone rf response are presented. The device parameters that control device performance are analyzed. These results demonstrate the viability of the IOSA for real time spectrum analysis of pulsed and cw rf signals.

Improvements of rf bandwidth resolution can be obtained by the use of larger collimated optical beams. This requires larger optical lens elements, and hence, larger crystals.

REFERENCES

1. M.K. Hamilton, D.A. Willie, and W.J. Miceli, "An Integrated Optical RF spectrum Analyzer," Proc. IEEE 1976 Ultrasonic Symp. Annapolis, MD.

2. M.K. Barnoski, B. Chen, T.R. Joseph, J.Y. Lee, and O.G. Ramer, "An Integrated Optic Spectrum Analyzer," IEEE Trans. on Circuits and Systems, CAS-26, 1113-1124 (Dec, 1979).

3. R.D. Weglein, M.E. Pedinoff, and H. Winston, "Diffraction Spreading of Surface Wave on $LiNbO_3$," Electron Lett. 6, 20 (1970).

4. R.V. Schmidt and I.T. Kaminow, "Metal-Diffused Optical Waveguide in $LiNbO_3$," Appl. Phys. Lett. 25, 458-460 (1974).

5. D. Kassai and E. Marom, "Aberration-Corrected Rounded-Edge Geodesic Lenses," J. Opt. Soc. Am. 69, 1242-1247 (1979).

6. B. Chen, E. Marom, and R.J. Morrison, "Diffraction Limited Geodesic Lens for Integrated Optic Circuits," Appl. Phys. Lett. 33, 511-513 (Sept., 1978).

7. B. Chen and O.G. Ramer, "Diffraction Limited Geodesic Lens for Integrated Optic Circuits," IEEE J. Quantum Electron, QE-15, 853-860 (1979).

8. G.E. Betts, J.C. Bradley, G.E. Marx, D.C. Schakest, and H.A. Trenchard, "Axially Symmetric Geodesic Lenses," Appl. Opt. 17, 2346-2351 (Aug., 1978).

9. T. Szabo and A. Slobodnik, "The Effect of Diffraction in the Design of Acoustic Surface Wave Devices," IEEE Trans. on Sonics and Ultrasonics SU-20, No. 3 (July, 1973).

10. I.C. Chang, "Acoustooptic Devices and Applications" IEEE Trans. on Sonics and Ultrasonics SU-23, 2, (January, 1976).

Table 1. Parameters of Collimation and Fourier Transform Lenses.

Focal length	18.8 mm
Diameter	8.7 mm
Depth at center	1.1 mm
Useful aperture	7.4 mm
f/no.	2.54

Table 2. Measured Attenuation Values for Titanium Diffusion $LiNbO_3$, obtained at 600 MHz and 1000 MHz.

Sample Number	Oxidation Time at 650°C, hr	Diffusion Time at 950°C, hr	Attenuation, dB/µsec
			600 MHz
N-90	0	5	1.9
N-91	0	8	7.0
N-92	2	5	0.54
N-95	4	6	1.9
Untreated surface (theoretical)	0	0	0.43

Table 3. Measured Velocity Dispersion Coefficier
for Ti:$LiNbO_3$ Substrates. Data Was Taken Over
the 400 to 800 MHz Range.

Sample	Oxidation Time at 600°C, hr	Diffusion Time at 1000°C, hr	$d\tau = \frac{1}{\tau}\left(\frac{d\tau}{df}\right)$, 10^{-5} MHz or $\frac{-1}{V_s}\left(\frac{dV_s}{df}\right)$
A	4	5	-1.59
B	0	5	-1.78
C	4	5	-1.57
D	0	5	-1.59
E	4	6	-1.55
F	4	6	-1.45
G	4	6	-1.57
H	4	6	-1.75

Table 4. Noise Characteristics of Surface-Channel
and Buried-Channel CCDs.

Operation	Process	Surface Channel	Buried Channel
Input	Fat zero input	178	0
Integration and transfer	Fast interface states	834	0
	Bulk states	∿0	90
	Dark current	<100	<100
Output	Reset amplifier	147	147
	MOSFET	68	68
Total noise electrons (rms)	All	873	210

Table 5. Charge Transfer Characteristics.

Carrier Mobility, cm^2/V-sec	Electrode Length, μm	t_{th}, sec	t_{tr}, sec	t_4, sec	f, MHz
600	10	$2.56x10^{-8}$	$9.67x10^{-9}$	$6.6x10^{-8}$	7.5
600	7.5	$1.44x10^{-8}$	$3.04x10^{-9}$	$3.35x10^{-8}$	14.9
600	6	$9.23x10^{-9}$	$1.29x10^{-9}$	$2.04x10^{-8}$	24.5
1000	10	$1.54x10^{-8}$	$5.80x10^{-9}$	$3.97x10^{-8}$	12.5
1000	7.5	$8.65x10^{-9}$	$1.83x10^{-9}$	$2.01x10^{-8}$	24.8
1000	6	$5.54x10^{-9}$	$7.76x10^{-10}$	$1.23x10^{-8}$	40.6

Table 6. Scaling of Laser Power and CCD Clocking to Maintain Performance.

Laser Power Output, mW	Total In/Out Insertion Loss, dB	Output Zeroth Order Power, μW	RF Power at Saturation, mW	Integration Time, μsec
1-5	17	30	200	19.2
15	17	300	200	19.2
15	11	1200	200	4.8
15	11	1200	1000	0.96

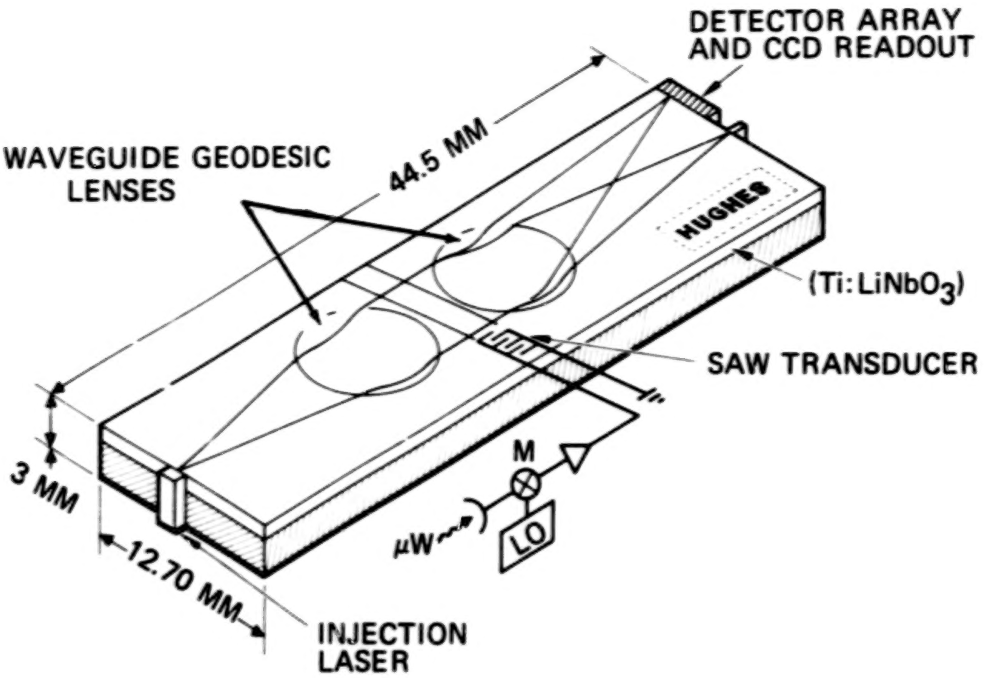

DETECTOR ARRAY
AND CCD READOUT
WAVEGUIDE GEODESIC
LENSES
44.5 MM
HUGHES
(Ti:LiNbO$_3$)
SAW TRANSDUCER
3 MM
M
μW
LO
12.70 MM
INJECTION
LASER

Figure 1.- Schematic of an integrated-optic spectrum analyzer.

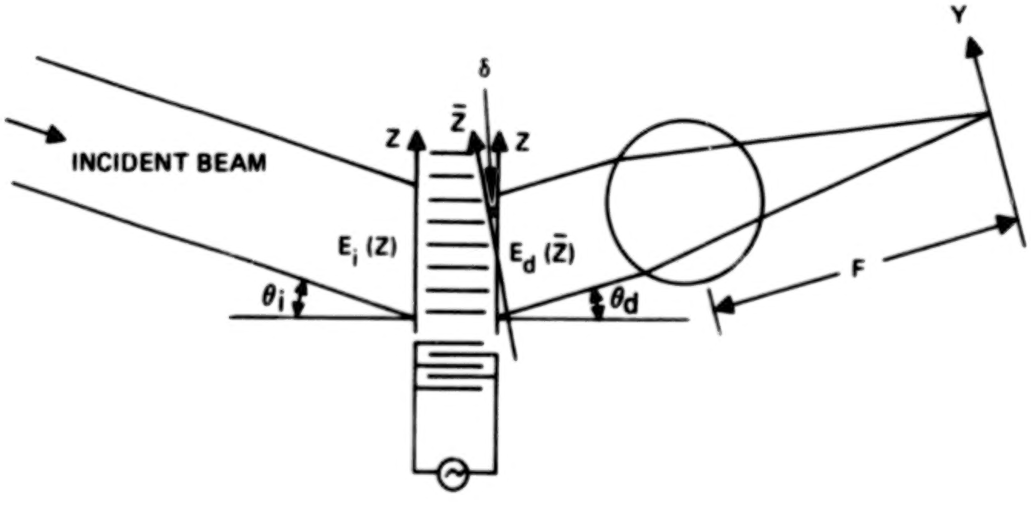

Y
δ
$\bar{Z}$
Z
Z
INCIDENT BEAM
E_i (Z)
E_d ($\bar{Z}$)
θ_i
θ_d
F

Figure 2.- Optical wavefront propagation through the acousto-optic interaction region and the transform lens.

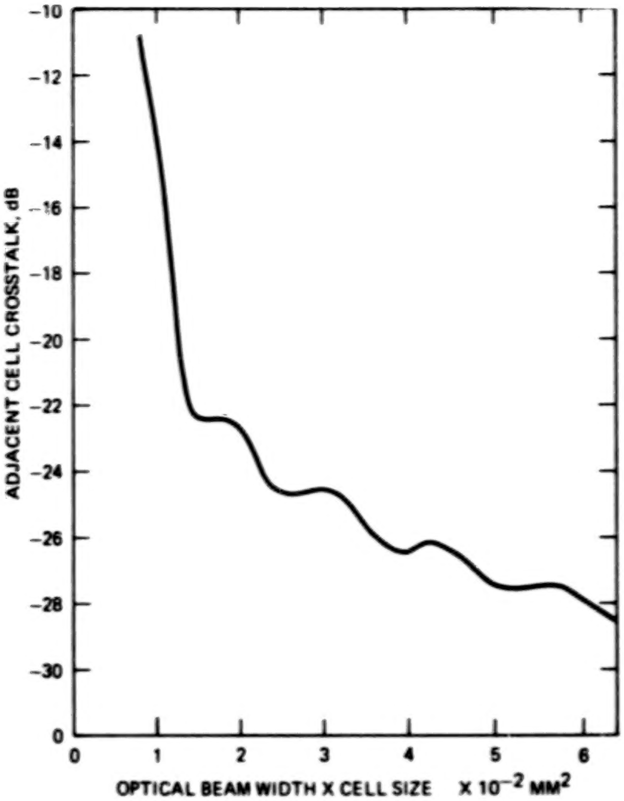


Figure 3.- Optical crosstalk in the adjacent cell of the detector array as a function of optical beam width and detector cell size.

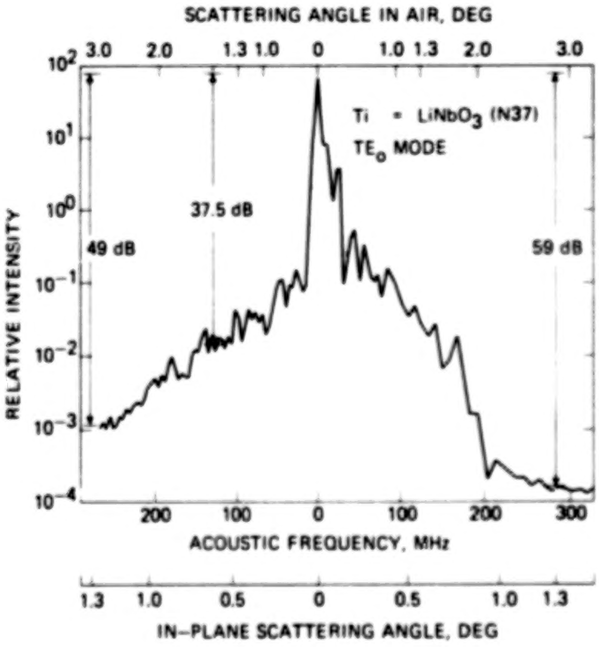


Figure 4.- In-plane scattering measurement of Ti:LiNbO$_3$ waveguide.

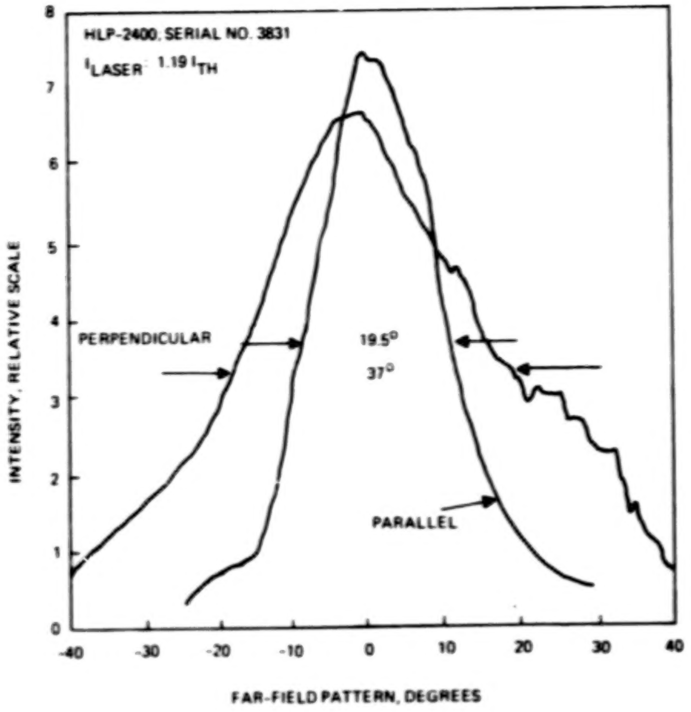


Figure 5.- Far field intensity pattern for a Hitachi HLP-2400 laser, Serial No. 3831.

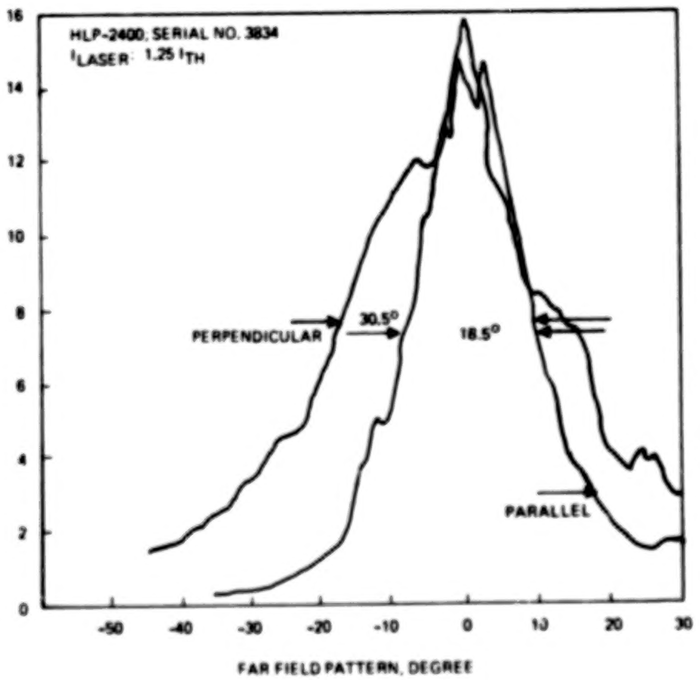


Figure 6.- Far field intensity pattern for a Hitachi HLP-2400 laser, Serial No. 3834.

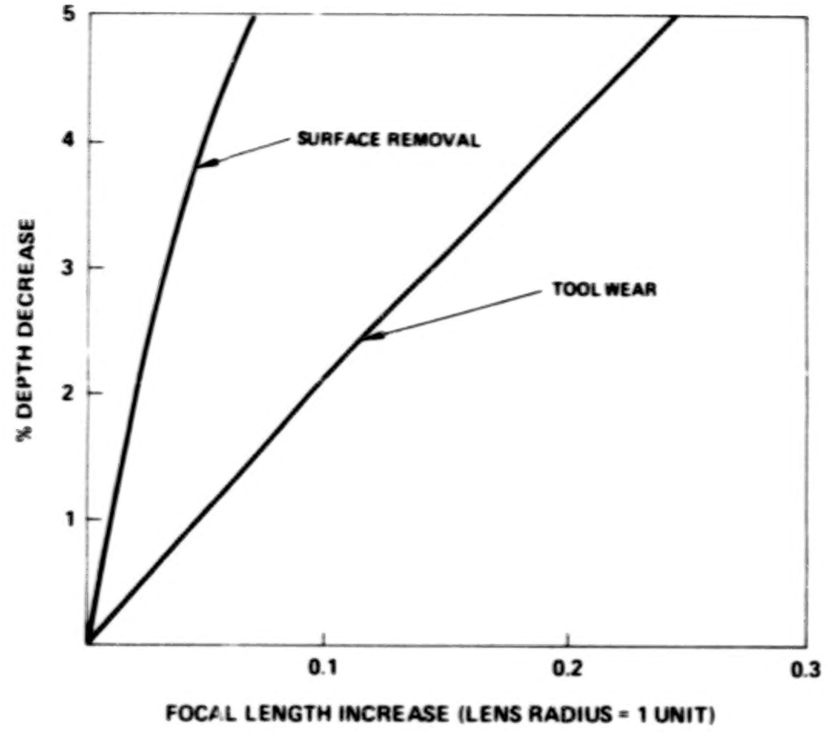


Figure 7.- Focal length increase versus percentage of depth decrease for both tool wear and surface removal.

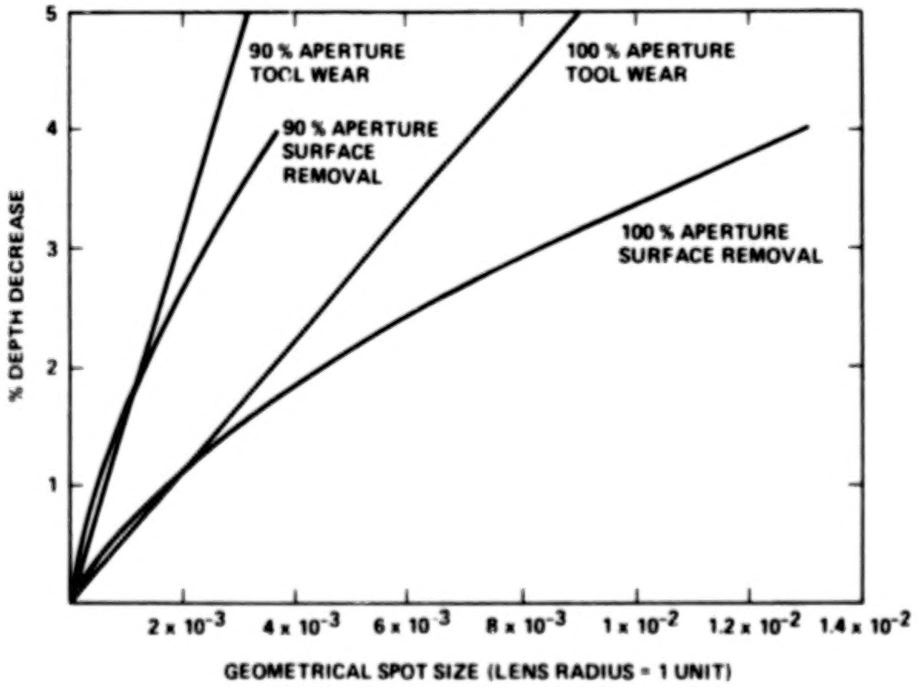


Figure 8.- Geometrical spot size versus percentage of depth decrease for both tool wear and surface removal. Two curves are provided for each one corresponding to 100% and 90% of the original useful aperture.

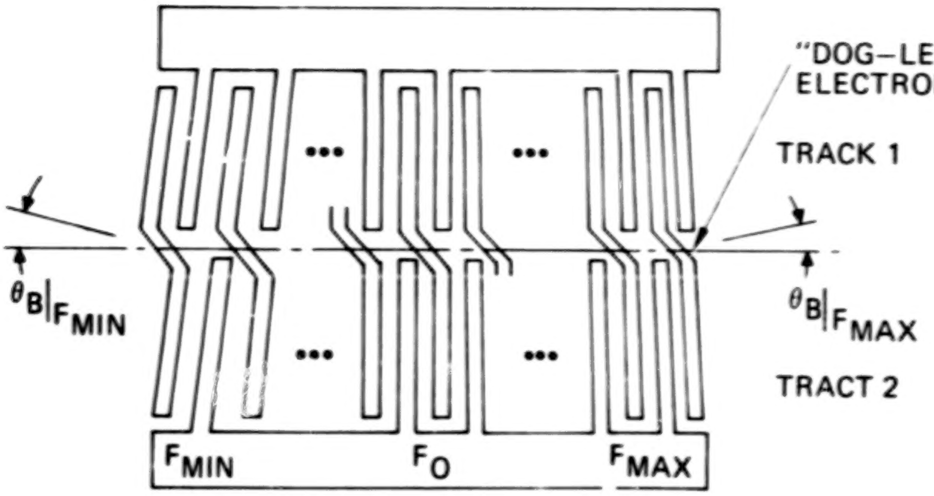

"DOG-LEG"
ELECTRODE
TRACK 1
$\theta_B|_{F_{MIN}}$
$\theta_B|_{F_{MAX}}$
TRACT 2
F_{MIN}
F_O
F_{MAX}

Figure 9.- Schematic of a modified chirp transducer with a two-track dog-leg electrode geometry.

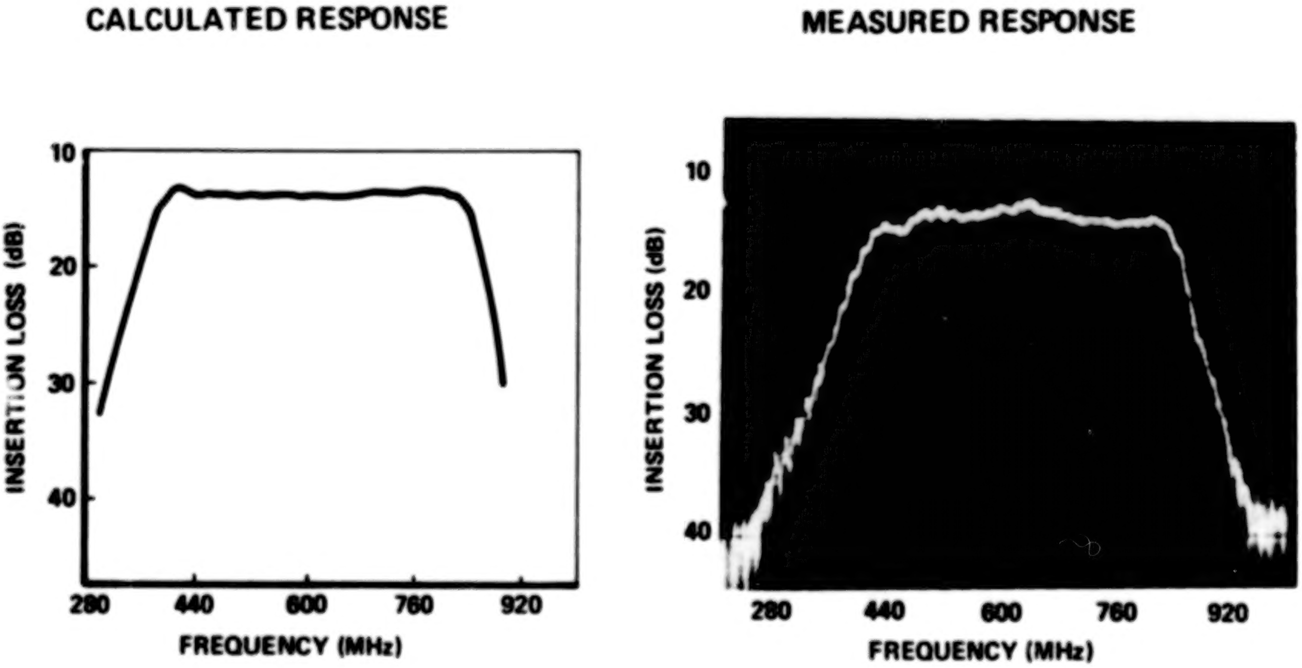


Figure 10.- Comparison between the calculated and measured responses of the 600-MHz modified chirp transducer.

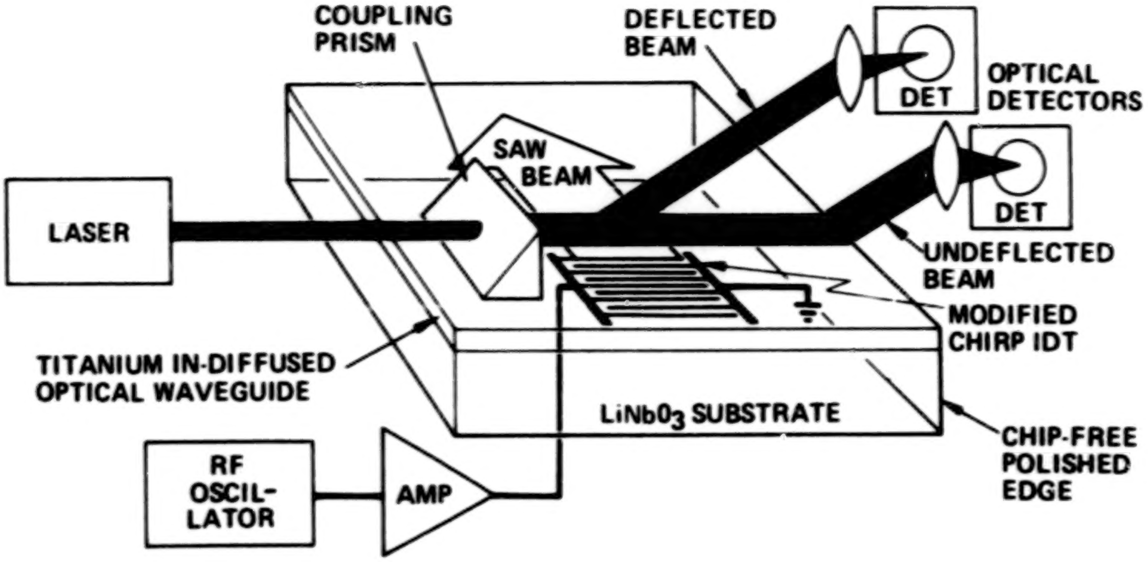
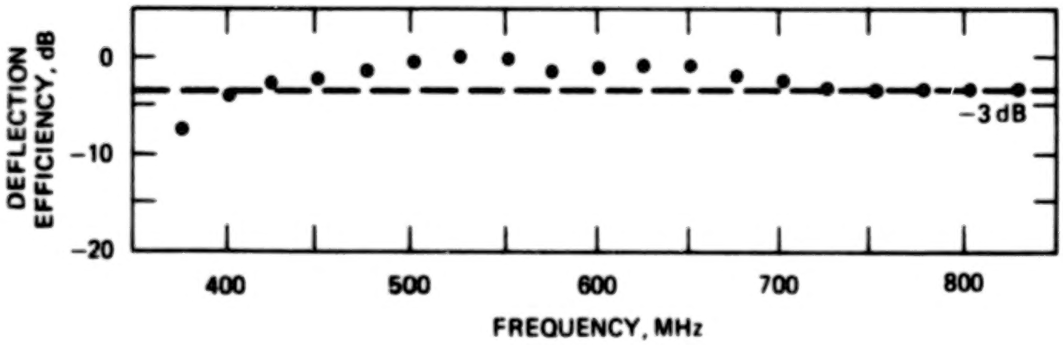

COUPLING PRISM
DEFLECTED BEAM
DET
OPTICAL DETECTORS
SAW BEAM
LASER
DET
UNDEFLECTED BEAM
MODIFIED CHIRP IDT
TITANIUM IN-DIFFUSED OPTICAL WAVEGUIDE
LiNbO3 SUBSTRATE
CHIP-FREE POLISHED EDGE
RF OSCIL-LATOR
AMP

Figure 11.- Test configuration used in the measurement of the Bragg acousto-optic deflection coefficient.



λ_{op} = 0.83 μm TE MODE
POWER LEVEL = 100 mW
0 dB = 1.6%
$BW_{3\,dB}$ = 420 MHz

Figure 12.- Measured acousto-optic deflection efficiency for a TE-mode GaAlAs laser.

BEST DOCUMENT AVAILABLE

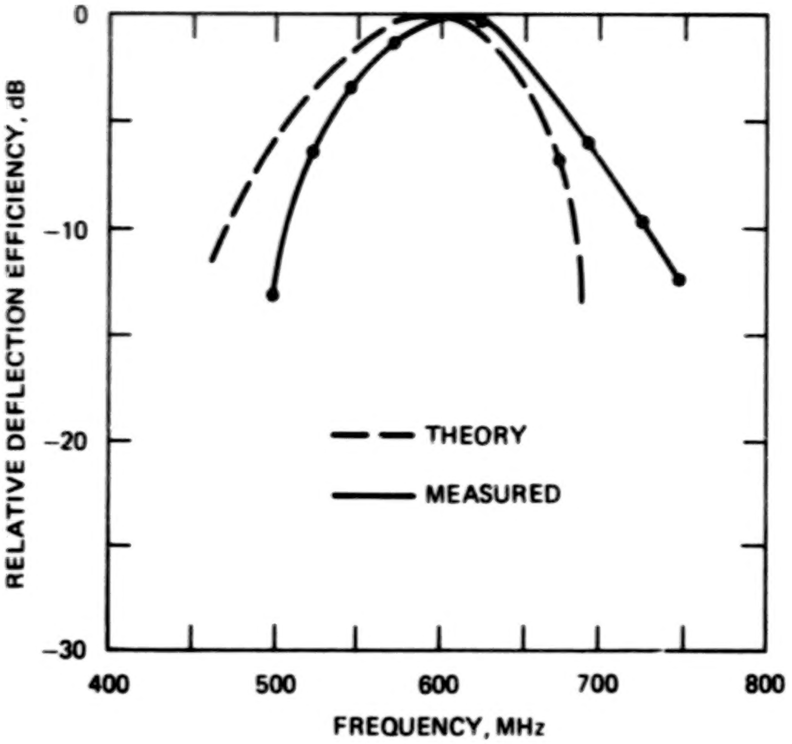
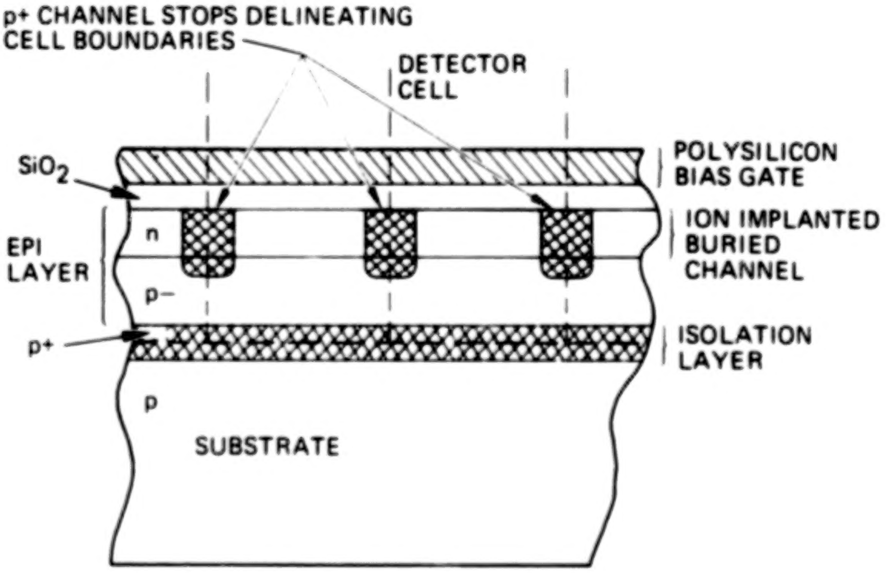


Figure 13.- Comparison of the theoretical and measured acousto-optic deflection efficiency when the sign of the electrode tilt is reversed.



p+ CHANNEL STOPS DELINEATING CELL BOUNDARIES
DETECTOR CELL
POLYSILICON BIAS GATE
SiO_2
EPI LAYER
n
p−
ION IMPLANTED BURIED CHANNEL
p+
ISOLATION LAYER
p
SUBSTRATE

Figure 14.- Cross section of buried p-channel MOS photodetector.

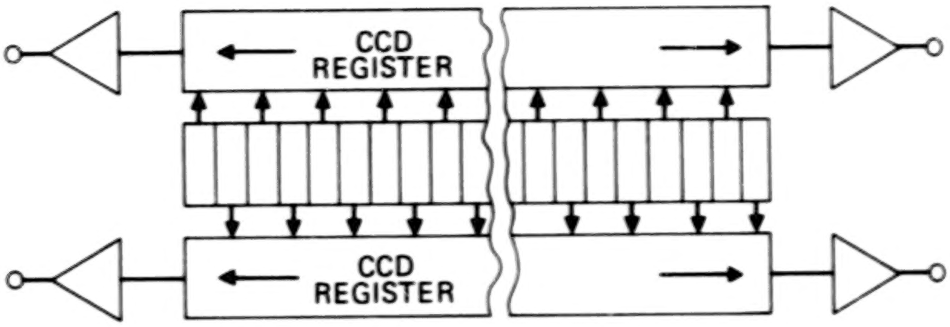

CCD REGISTER
CCD REGISTER

Figure 15.- Schematic showing detector array, CCD output register, and MOS output amplifiers. Arrows denote direction of charge transfer.

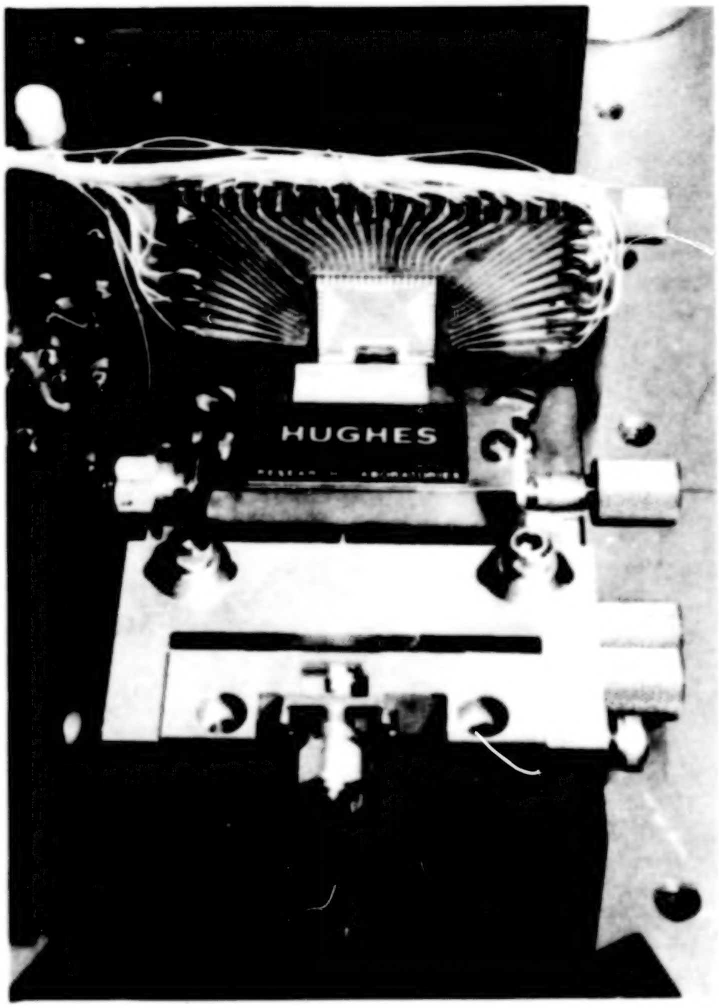


Figure 16.- Close-up photoghaph of the assembled IOSA.

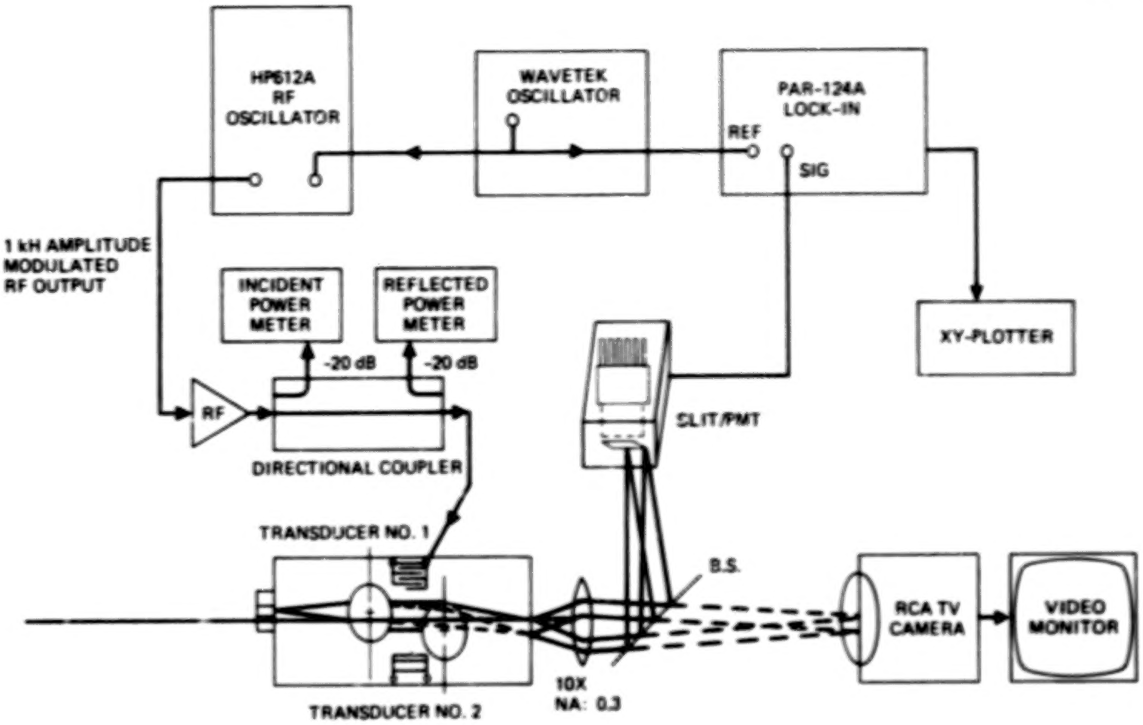

HP612A
RF
OSCILLATOR
WAVETEK
OSCILLATOR
PAR-124A
LOCK-IN
REF
SIG
1 kH AMPLITUDE
MODULATED
RF OUTPUT
INCIDENT
POWER
METER
REFLECTED
POWER
METER
-20 dB
-20 dB
XY-PLOTTER
RF
DIRECTIONAL COUPLER
SLIT/PMT
TRANSDUCER NO. 1
B.S.
RCA TV
CAMERA
VIDEO
MONITOR
TRANSDUCER NO. 2
10X
NA: 0.3

Figure 17.- Schematic of arrangement used to measure the acousto-optic response of the IOSA.

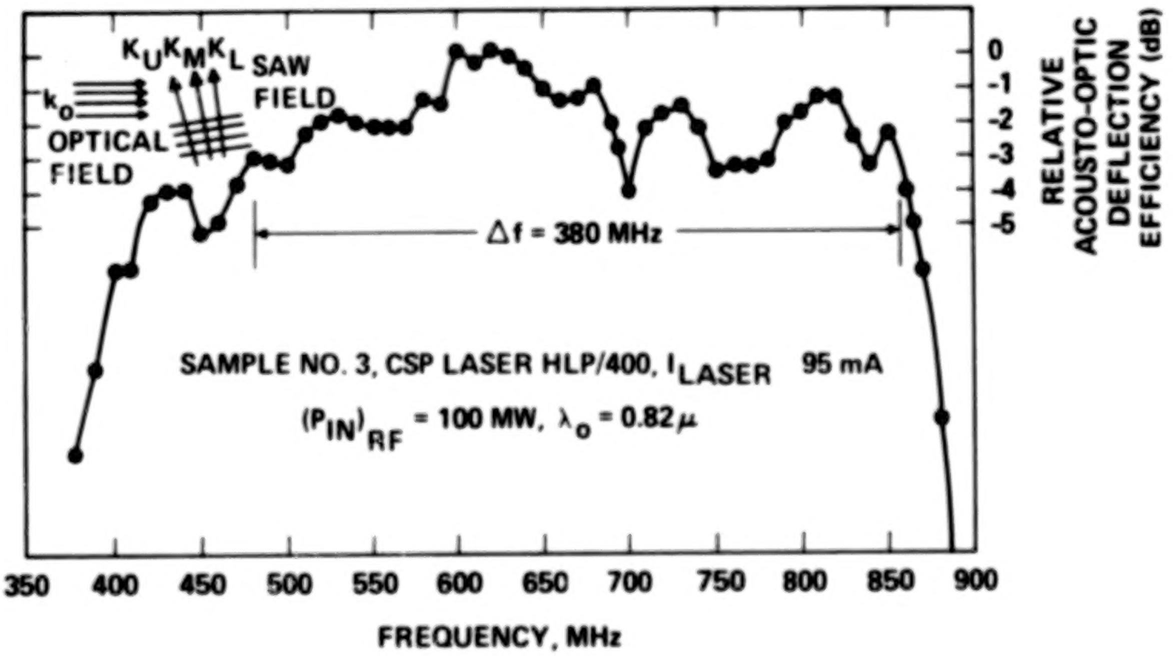


Figure 18.- Acousto-optic response measured on a device with a Hitachi HLP 1400 CSP laser.

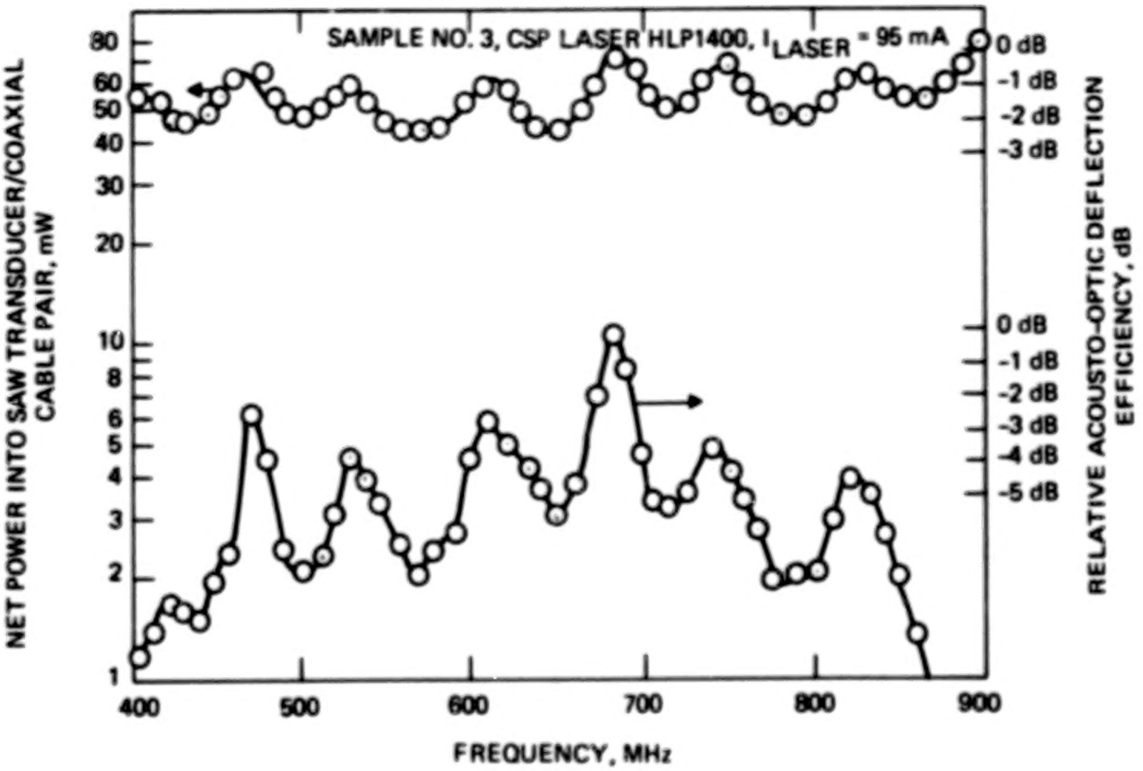
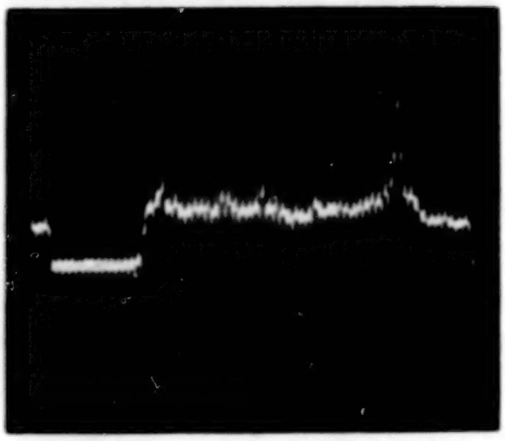


Figure 19.- Acousto-optic response and net rf power into SAW versus frequency. The clear-cut correspondence in the position of the two plots leads to the conclusion that the irregular structure in the acousto-optic response is due to rf matching irregularities.



DETECTOR/CCD OUTPUT FOR A 725 MHz/100 mW
RF SIGNAL APPLIED TO THE IOSA

HORIZONTAL: 20 μs/DIV
VERTICAL: 50 mV/DIV
3 dB RESOLUTION: 8 MHz

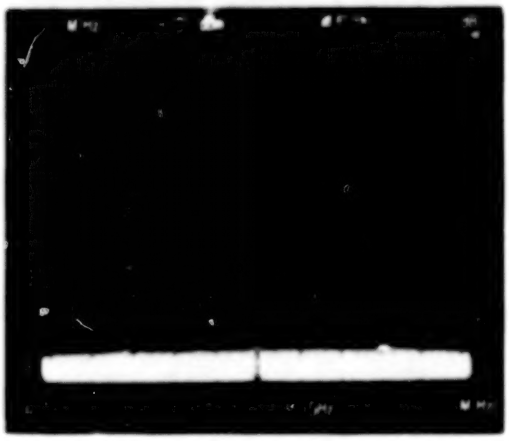


Figure 20.- Single-tone frequency resolution on the IOSA device. Sample 4.

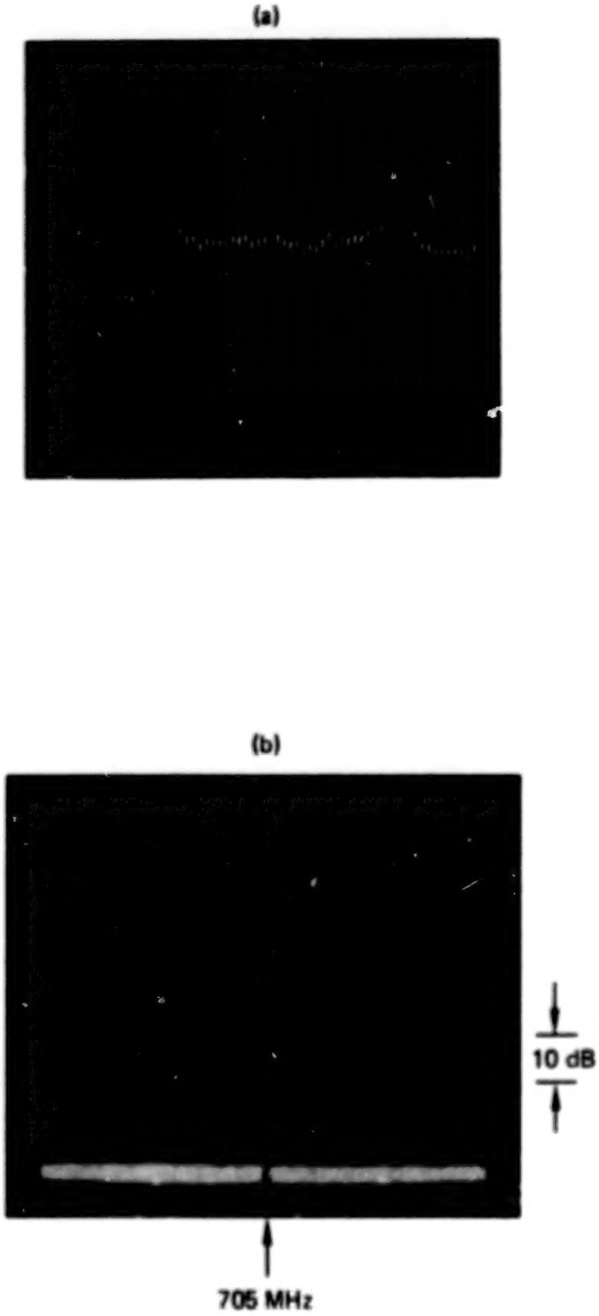


Figure 21.- Response from one half of the detector/CCD array when driven by a 100 mW, 705 MHz rf signal. (Vert: 50 mV/div, Horiz: 20 μsec/div).

BEST DOCUMENT AVAILABLE

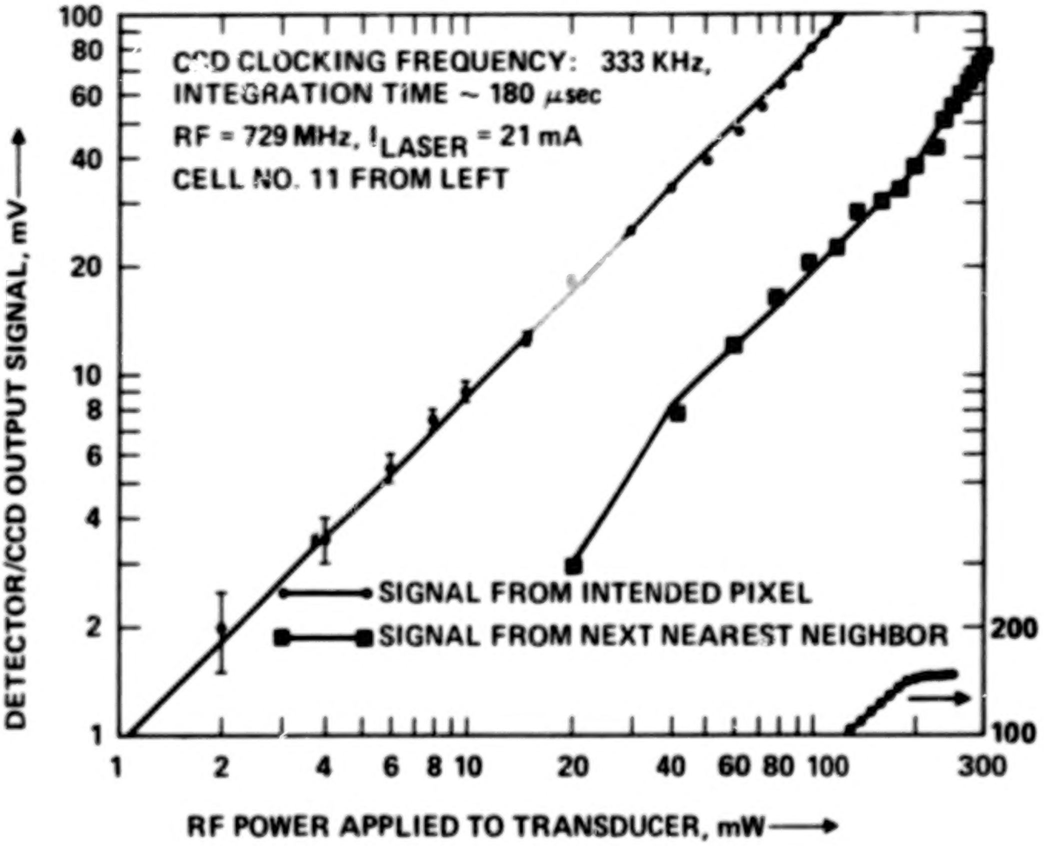
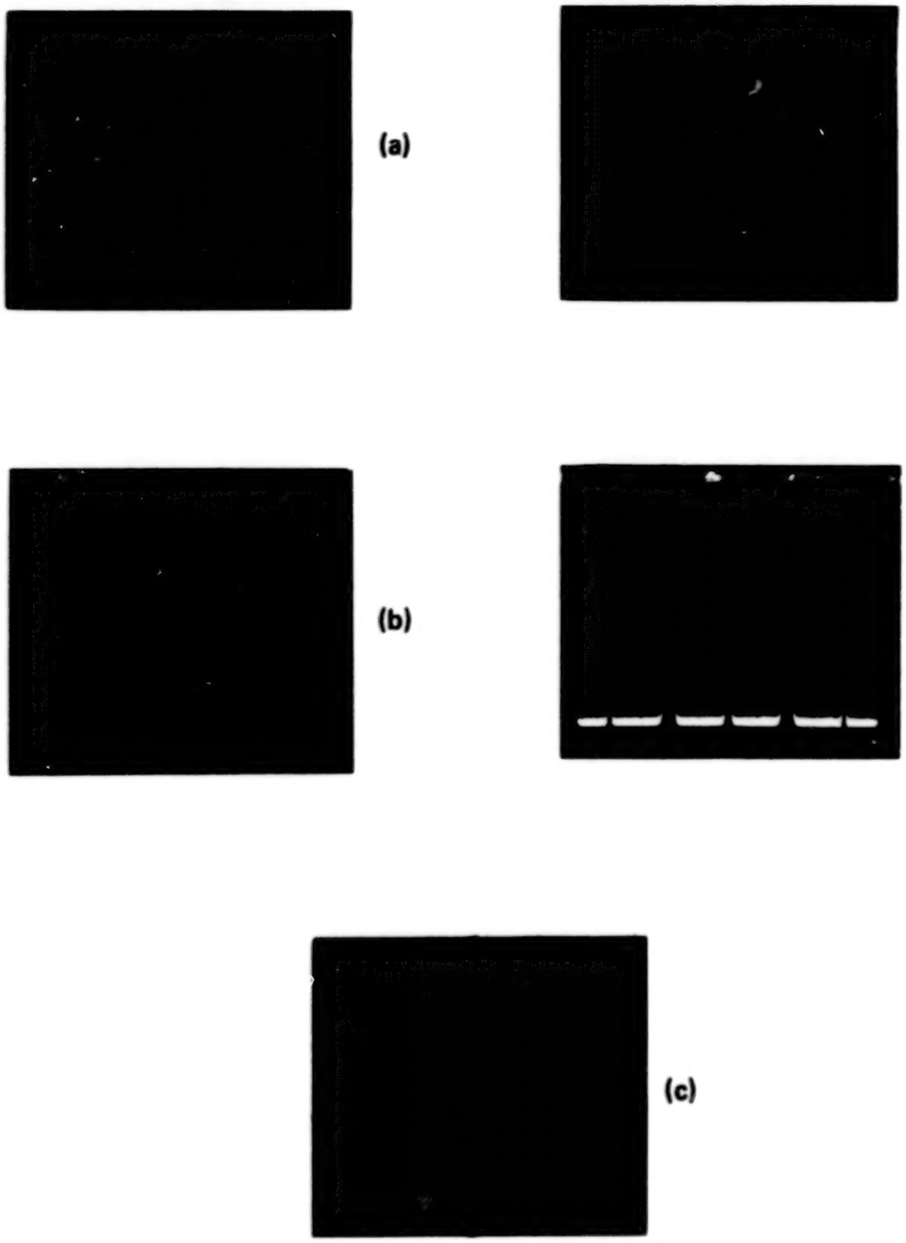
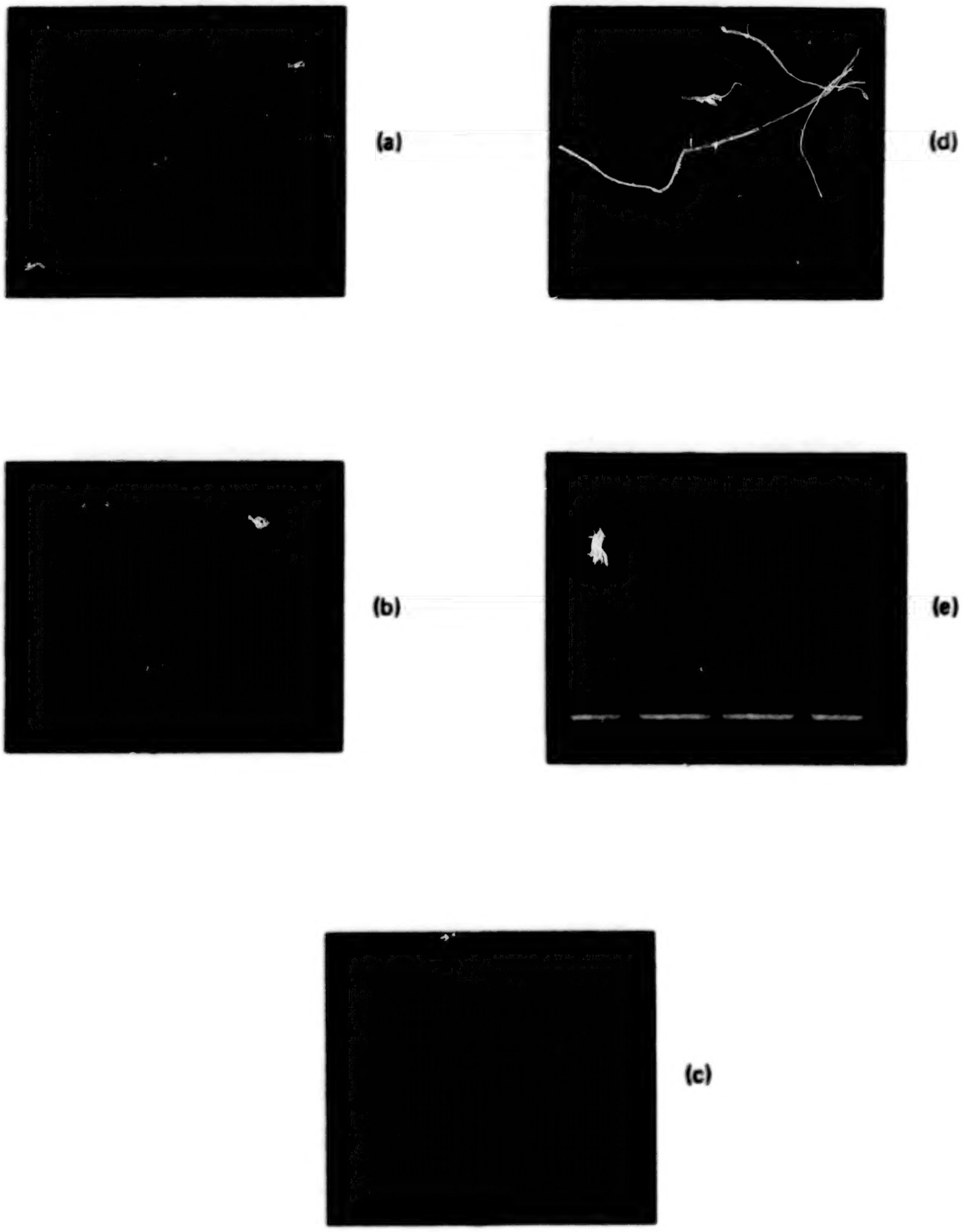


Figure 22.- Dynamic range and cross talk measured at a CCD clocking frequency of 333 kHz (Sample 4, HLP 2400 laser, pixell 11 from the left).



VERTICAL SCALE: 50 mV/div FOR a), b), c) AND d). a) AND c) COMPOSITE 100 ELEMENT DISPLAY. HORIZONTAL SCALE: 20 μsec/div AND 10 μsec/div. b) AND d) SEPARATE DISPLAY OF THE TWIN DETECTOR ARRAY CHANNELS. HORIZONTAL SCALE: 20 μsec/div AND 10 μsec/div

Figure 23.- Two-tone response on Sample 4, $\Delta F = 8$ MHz, $F_o = 689$ MHz, $(P_{rf})_{tot} = 200$ mW, HLP 2400 laser, $I_{laser} = 21$ mA.



VERTICAL SCALE: 50 mV/div FOR a), b), c) AND d). a) AND c) COMPOSITE 100 ELEMENT DISPLAY. HORIZONTAL SCALE: 20 μsec/div AND 10 μsec/div. b) AND d) SEPARATE DISPLAY OF THE TWIN DETECTOR ARRAY CHANNELS. HORIZONTAL SCALE: 20 μsec/div AND 10 μsec/div

Figure 24.- Two-tone response on Sample 4, $\Delta F=12$ MHz, $F_o=689$ MHz, $(P_{rf})_{tot}=200$ mW, HLP 2400 laser, $I_{laser}=21$ mA.

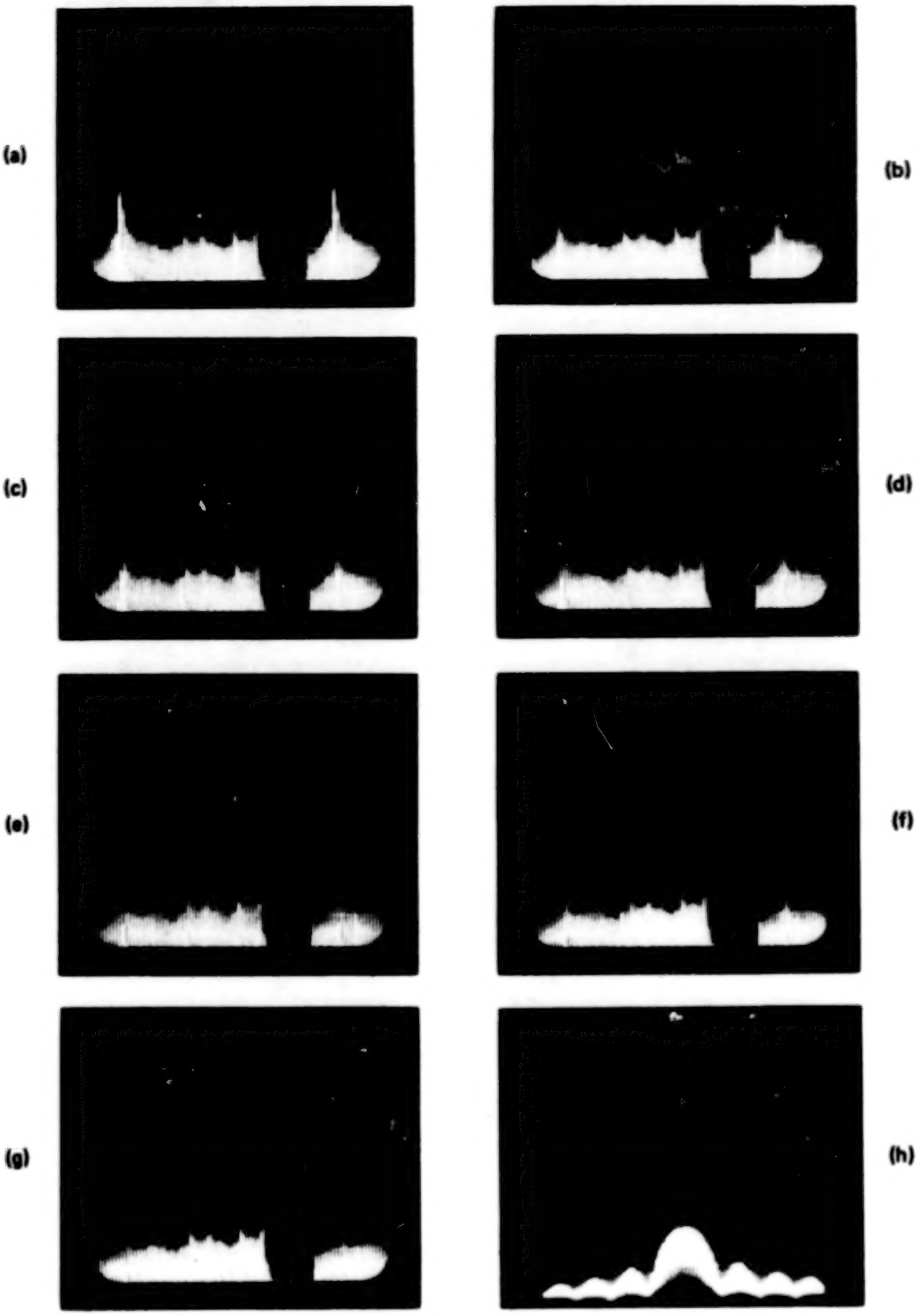


Figure 25.- RF pulse measurements. CCD clocking frequency 1.67 MHz. I_{laser}=27 mA, Vert scale: 50 mV/div, Horz. 5 μsec/div. a) cw rf power 100 mW. Pulse duty factors: b) 0.5, c) 0.4, d) 0.3, e) 0.2, f) 0.1 and g) 0. h) display the rf spectrum of a pulse train with a duty factor 0.1.

APPLICATIONS OF ELECTRO-OPTIC GRATINGS IN INTEGRATED OPTICAL SIGNAL PROCESSING DEVICES*

Carl M. Verber
Battelle Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

SUMMARY

The electro-optic grating is an easily fabricated component which allows a rapid and efficient interaction with an optical wave in a planar electro-optic waveguide. The operation of such gratings and their use as intensity modulators, spatial light modulators, and components in correlators and in a variety of computational units is described.

INTRODUCTION

The electro-optic effect has been the most important interaction mechanism used in the design of the active integrated optical devices employing channel waveguides.[1] In these devices, the designer is able to take advantage of the lateral confinement of the light which allows long interaction lengths and consequently low voltages. In the present paper we will, in contrast, be concerned with the use of the electro-optic effect in planar integrated optical structures. In this planar geometry, the guided light is confined to a several micrometer-thick waveguide at the surface of a suitable electro-optic substrate. Lateral diffraction is not inhibited, but the confinement close to the surface allows the use of moderate voltages to obtain a strong interaction with the guided light. In particular we will show how the interaction of a relatively broad guided beam with an electro-optically induced thick phase (Bragg) grating[2] can be used to advantage in a variety of integrated optical devices. We begin by describing the properties of the gratings and then discuss a number of modulators, correlators and computational units which incorporate these gratings.

THE ELECTRO-OPTIC GRATING

The basic geometry of the I.O. grating electrodes is shown in figure 1. The electrodes are defined by standard photolithographic techniques upon the surface** of a planar electro-optic optical waveguide such as Ti-indiffused[4] $LiNbO_3$. It is conventional to make the electrode line-wide equal the spacing between adjacent fingers, so the required photolighographic resolution is $\Lambda/4$.

---

* Work supported in part by Air Force Office of Scientific Research.

** Often a sputtered glass buffer layer is employed[3] to isolate the guided wave from perturbation due to the metallization pattern.

The application of a voltage across the electrodes results in a periodic electric field which, via the electro-optic effect, gives rise to a periodic modulation of the index of refraction of the waveguide material and thus to the mode index of the guided wave. The guided wave sees this periodic perturbation as a thick phase grating and will be diffracted as indicated in figure 1 when incident upon the grating at the Bragg angle, defined by

$$\sin\theta_B = \frac{\lambda}{2\Lambda} \tag{1}$$

where λ is the optical wavelength in the medium.

The diffraction efficiency[5] is

$$\eta = \sin^2 \frac{\pi\Delta n d}{\lambda_o \cos\theta_B} \tag{2}$$

where Δn is the amplitude of the periodic index modulation. The magnitude of Δn is determined by the product of the applied field strength E, and the appropriate electro-optic coefficient r_{ij} according to[6]

$$\Delta n = \frac{1}{2} n^3 r_{ij} E \tag{3}$$

where n is the average index of refraction. Inclusion of geometric effects[7] results in an index modulation

$$\Delta n = \frac{1}{2} n^3 r_{ij} \frac{2}{\pi} \left(\frac{V}{\Lambda/4}\right) \tag{4}$$

or a diffraction efficiency of

$$\eta(V) = \sin^2 \left(\frac{4n^3 r_{ij} d}{\lambda_o \cos\theta_B} V\right) \tag{5}$$

Ignoring buffer layer effects, eq. (5) indicates that for Λ = 8 µm, a He-Ne laser and d = 2 mm, we get 100% diffraction efficiency for V = 3.1 volts. Since high diffraction efficiencies are readily achieved and the electrode capacitances are quite low, it is evident that the electro-optic Bragg effect can be utilized to make a high performance modulator. This device was originally suggested by Hammer and Phillips,[2] and has more recently been employed by Holman[8] to make a high performance planar modulation with a 69% optical throughput.

THE INTEGRATED OPTICAL SPATIAL LIGHT MODULATOR

The basic grating structure can be extended as shown in figure 2 by introducing electrodes which allow segments of the grating to be individually addressed. In this manner, one can impose a transverse amplitude modulation upon the diffracted beam. The undiffracted beam will of course have a complementary modulation. The grating structure is now operating as an electrically addressable integrated optical spatial light modulator (IOSLM) and can, in principle, be used to modulate an arbitrarily wide guided wave. The modulator can be used in an analog or a binary mode, although there will obviously be a finite number of addressable segments. The largest such IOSLM we have fabricated thus far[9] is composed of 32 segments 200 μm wi [illegible] spans a 6.4 mm-wide guided wave.

It should be possible to reduce the width of each grating segment to 50 μm or less. It would then be possible to use the IOSLM in the signal and filter planes of a planar optical Fourier transform device.[10] However, even with the larger grating segment size, several useful functions can be performed by incorporating a SAW transducer on the same substrate as the IOSLM. If the orientations of the EO and acoustic gratings are chosen so that only doubly diffracted light is detected, then by introducing a single acoustic pulse which is shorter than a single IOSLM element, the transverse modulation produced by the IOSLM is converted to a temporal modulation of the output (doubly diffracted) beam. Parallel-to-serial conversion is thus accomplished. Another application involving the combined IOSLM SAW structure is the correlator, which is shown schematically in figure 3. Such a device has been constructed[9] and operates on a 32 bit word at a data-rate of 17.5 Mbit/sec.

COMPUTATION WITH GRATINGS

Consider once again the simple grating structure shown in figure 1. If voltage V_A is applied to one electrode and voltage V_B to the other, the diffracted light intensity is given by

$$I = I_o \sin^2 [a(V_A - V_B)] \tag{6}$$

where a is a constant. For $I \lesssim 0.1\, I_o$ we have

$$I = I_o\, a^2 (V_A - V_B)^2 \tag{7}$$

Now arrange N such electrode sets in a vertical line so that they can each be addressed by the same broad optical beam. If the N components of the vector $\vec{A}$ are applied to one electrode of each set and the corresponding component of the vector $\vec{B}$ is applied to the opposed electrodes then the total deflected intensity is

$$I = \sum_{i=1} I_i = \sum_{i=1} a^2 (V_{Ai} - V_{Bi})^2 \tag{8}$$

We have therefore calculated the magnitude of the difference of the two vectors.

Vector multiplication can also be performed as is suggested in figure 4. In this structure the diffracted intensity from the ith segment is

$$I_i = I_o\, \eta_{Ai}\, \eta_{Bi} \tag{9}$$

Of course, each η is actually proportional to the square of the voltage difference but it can be shown that by a simple signal processing technique, the desired $A_i B_i$ term can be extracted from I_i. The structure shown in figure 4 is therefore capable of performing the scalar product of two vectors.

It should also be noted that if binary signals are applied to the grating structures then standard logic operations may be performed. The structure used for subtraction can be used to generate the EXCLUSIVE OR operation, or the NOT operation, while the structure used for multiplication can be used for the logical AND. These operations may also be carried out in parallel, and in the case of binary signals, the complications arising from the the $\sin^2$ response are no longer present.

SUMMARY

We have shown a variety of applications of electro-optically induced Bragg gratings in integrated optical signal processing and computation devices. The gratings are easy to fabricate and operate efficiently on relatively low voltages. The design principles are well known and reliable. It is therefore quite likely that a large number of additional devices employing similar grating structures will be developed.

REFERENCES

1. Alferness, R. C.: Guided Wave Devices for Optical Communication. IEEE J. of Quant. Electr., vol. QE-17, no. 6, June 1981, pp. 946-959.

2. Hammer, J. M. and Phillips, W.: Low-Loss Single-Mode Optical Waveguides and Efficient High-Speed Modulators of $LiNb_xTa_{1-x}O_3$ on $LiTaO_3$. Appl. Phys. Lett., vol. 24, no. 11, June 1974, pp. 545-547.

3. Tangonan, G. L., Persechini, D. L., Lotspeich, J. F., and Barnoski, M. K.: Electrooptic Diffraction Modulation in Ti-diffused $LiTaO_3$. Appl. Opt., vol. 17, no. 20, Oct. 1978, pp. 3259-3263.

4. Schmidt, R. V. and Kaminow, I. P.: Metal Diffused Optical Waveguides in $LiNbO_3$. Appl. Phys. Lett., vol. 25, no. 8, Oct. 1974, pp 458-460.

5. Kogelnik, H.: Coupled Wave Theory for Thick Hologram Gratings. Bell Syst. Tech. J., vol. 48, no. 9, Nov. 1969, pp. 2909-2947.

6. Kaminow, I. P.: Introduction to Electrooptic Devices. Academic Press, New York, 1974, 409 pp.

7. Kaminow, I. P. and Stulz, L. W.: A Planar Electrooptic-Prism Switch, IEEE J. of Quant. Electr., vol. QE-11, no. 8, August 1975, pp. 633-635.

8. Holman, R. L. and Cressman, P. J.: A Very High Throughput Damage-Resistant Lithium Niobate Waveguide Modulator. Technical Digest, Topical Meeting on Integrated and Guided-Wave Optics, Incline Village, NV, January 28-30 (1980), paper TuE6.

9. Verber, C. M., Kenan, R. P., and Busch, J. R.: Correlator Based on an Integrated Optical Spatial Light Modulator. Appl. Opt., vol. 20, no. 9, May 1981, pp. 1626-1629.

10. Verber, C. M., Kenan, R. P., and Busch, J. R.: An Integrated Optical Spatial Filter. Opt. Commun., vol. 34, no. 1, July 1980, pp. 32-34.

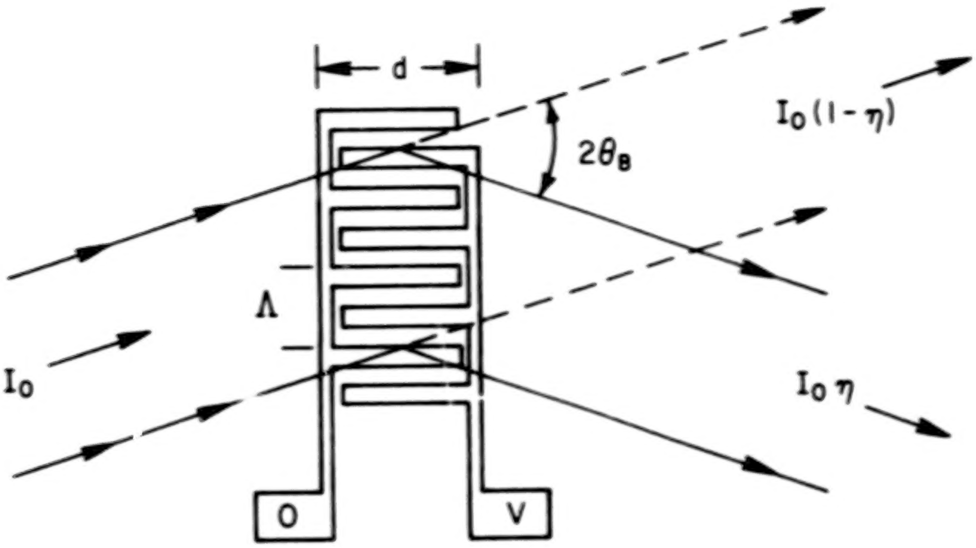

d
2θ_B
$I_0(1-\eta)$
Λ
I_0
$I_0\eta$
0
V

Figure 1.- Basic electrode structure. The grating wavelength Λ and depth d are indicated.

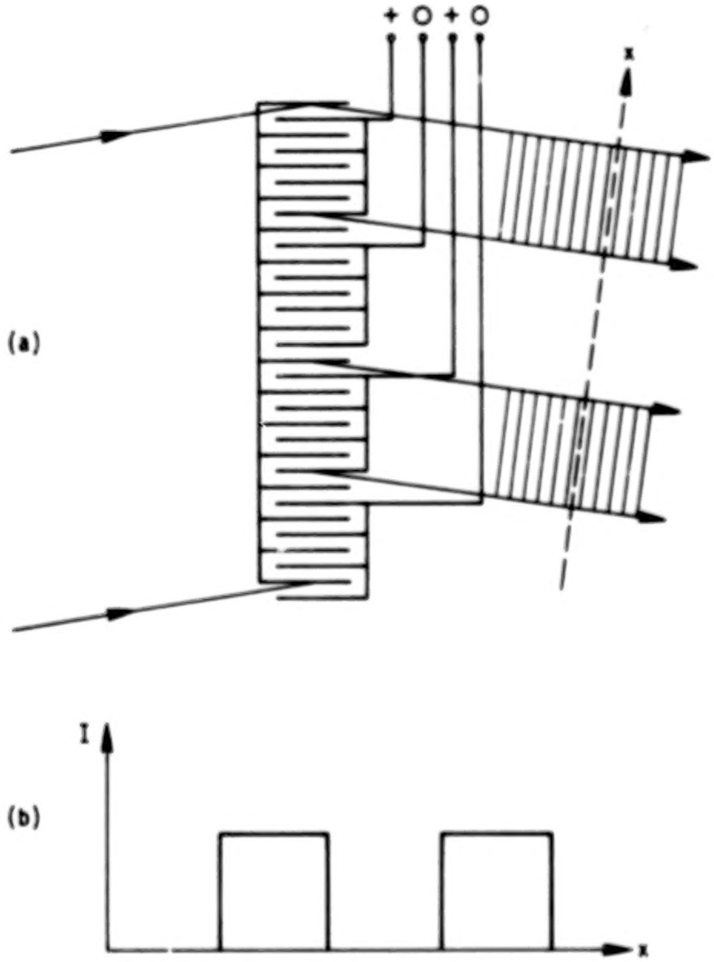


Figure 2.- The electrode structure is addressable in segments to form an integrated optical spatial light modulator. If the common electrode is at zero potential then the voltage pattern indicated in (a) produces the transverse amplitude modulation shown in (b).

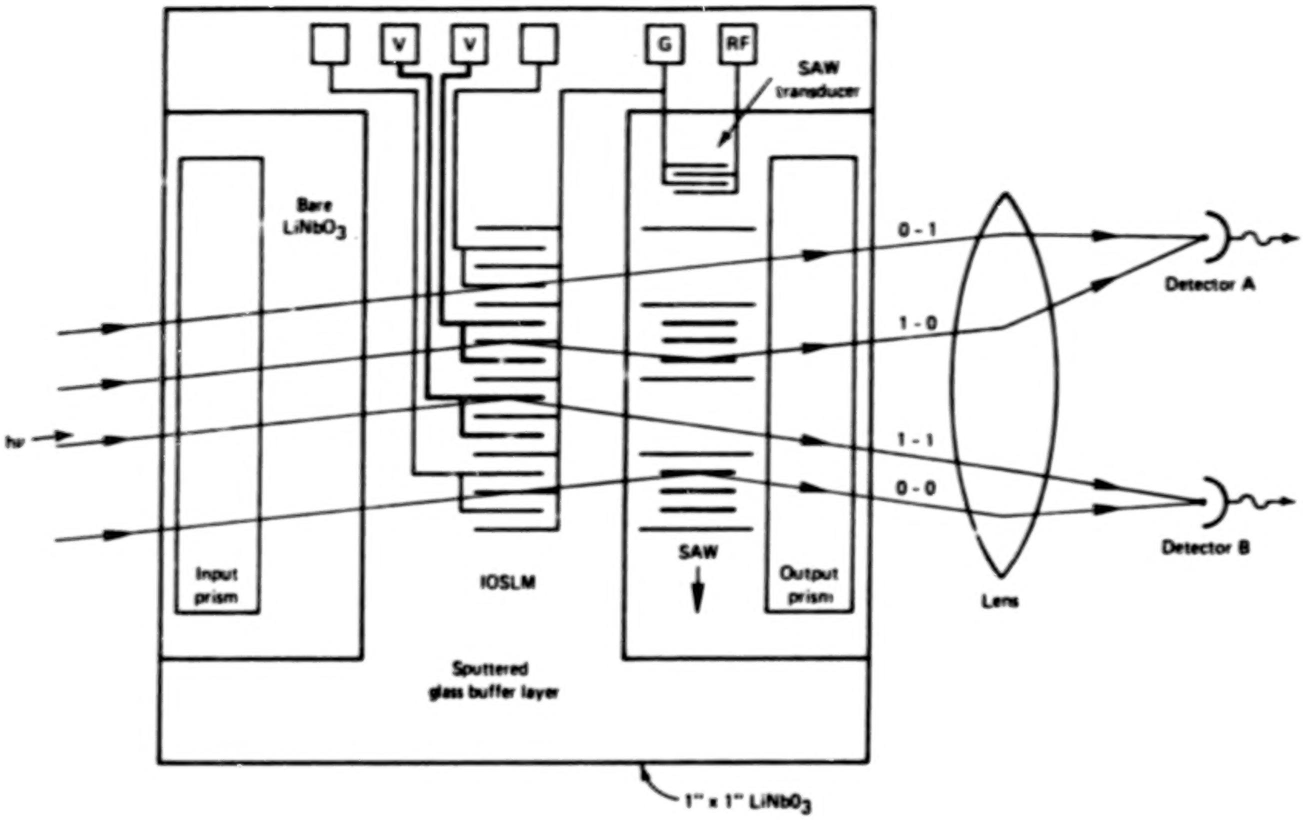

V
V
G
RF
SAW transducer
Bare LiNbO$_3$
0 - 1
1 - 0
1 - 1
0 - 0
Detector A
Detector B
hv
Input prism
IOSLM
SAW
Output prism
Lens
Sputtered glass buffer layer
1" x 1" LiNbO$_3$

Figure 3.- Correlator incorporating a programmable IOSLM and a SAW transducer with digital modulation.

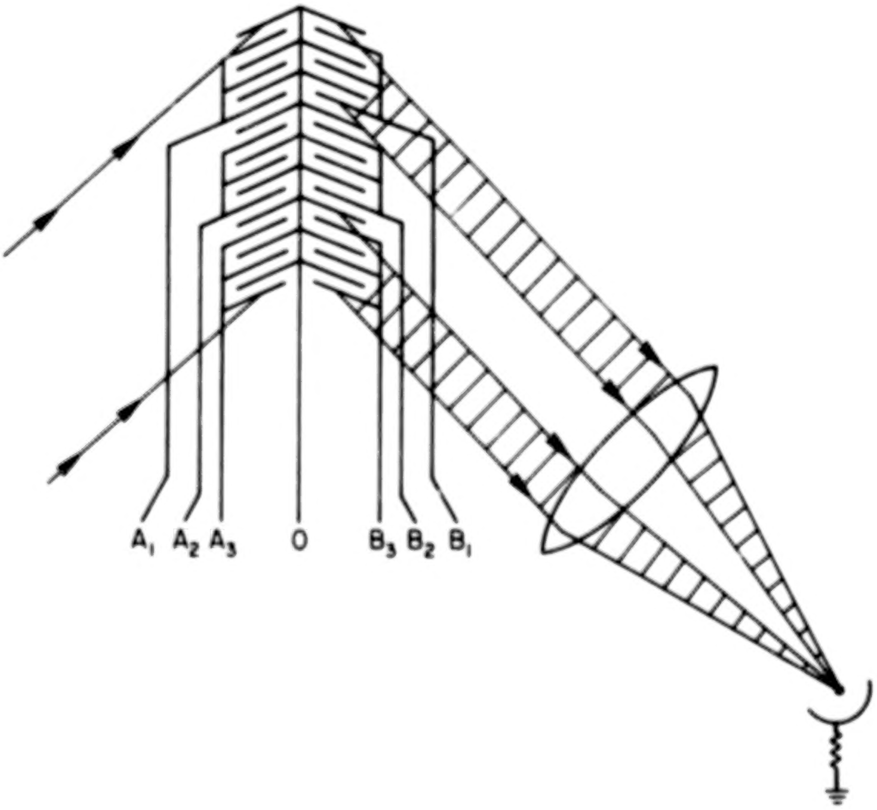

A$_1$
A$_2$
A$_3$
0
B$_3$
B$_2$
B$_1$

Figure 4.- Electrode structure for vector multiplication.

FOUR- AND FIVE-LAYER SILICON-CLAD DIELECTRIC WAVEGUIDES*

Glen McWright and T. E. Batchman
University of Virginia
Charlottesville, Virginia

ABSTRACT

Computer modeling studies on four-layer silicon-clad planar dielectric waveguides indicate that the attenuation (α) and mode index (β/K) behave as exponentially damped sinusoids as the silicon thickness is increased. The observed effect can be explained as a periodic coupling between the guided modes of the lossless structure and the lossy modes supported by the high-refractive index silicon. Furthermore, the attenuation and mode index are significantly altered by conductivity changes in the silicon. An amplitude modulator and phase modulator have been proposed using these results. Predicted high attenuations in the device may be reduced significantly with a silicon dioxide buffer layer.

INTRODUCTION

A need has arisen for direct optical modulation technology. This need has arisen from the search for faster digital switches, higher capacity data channels, and light, compact data preprocessing equipment for satellites. One promising technology that has been examined is the modulation of a guided light wave via photoconductivity changes in a semiconductor cladding.

Computer modeling studies on four- and five-layer, silicon-clad, planar dielectric waveguides indicate that the propagation characteristics can be altered by changes in the complex permittivity of the silicon and in the thickness of the silicon. Using these predictions, an intensity modulator and a phase modulator based on photon-induced conductivity changes in the semiconductor cladding have been studied.

SEMICONDUCTOR-CLAD WAVEGUIDES

The four-layer planar waveguide structure under consideration is shown in Figure 1, where it is assumed that light is propagating in the dielectric (N_3) and all materials are lossless except for the semiconductor (N_2). We desire to solve for the complex mode propagation constant ($\alpha + j\beta$).

---

*Research sponsored by NASA-Langley Research Center under Grant NSG-1567.

One technique (1) extends Maxwell's equations and boundary conditions to numerically solve a transcendental equation relating the attenuation constant (α) and phase constant (β) to the material types and thicknesses of the waveguide structure (hereafter refered to as PROGRAM WAVES).

A more efficient method (2) utilizes a matrix representation of Maxwell's equations, field solutions and boundary conditions in each waveguide layer (hereafter refered to as PROGRAM MODEIG). The matrices are multiplied and a characteristic matrix for the entire structure is obtained which yields the attenuation constant and phase constant.

The waveguide consists of a semi-infinite glass substrate, a polystyrene core of thickness 1 micrometer, a silicon cladding of .01 micrometer to 10 micrometers in thickness, and a semi-infinite layer of air. Each material is characterized by a complex relative permittivity, $\hat{\varepsilon}$; a free space wavelength of 632.8 nanometers is assumed and material parameters are shown for this wavelength (Fig. 1) Layers N_1, N_3, and N_4 are lossless dielectrics, so ε_r is real; however, at optical frequencies the permittivity of the silicon (N_2) is complex ($\varepsilon_r = \varepsilon'_r + j\varepsilon''_r$), and the complex part is a linear function of the conductivity ($\varepsilon''_r = \sigma/\omega\varepsilon_0$).

PREDICTED CHARACTERISTICS

The curves presented in Figures 2 and 3 were generated by repeated use of our own PROGRAM WAVES and later confirmed with PROGRAM MODEIG. The silicon cladding was varied from .01 micrometer to 10 micrometers and the complex mode propagation constant was calculated. The expected result was that as the cladding thickness was reduced to zero, the attenuation decreases to zero in a well-behaved manner; however, the results were not well-behaved when the silicon thickness falls below 1 micrometer. The curves are similar to exponentially damped sinusoids. Extreme β/K variations correspond to median values in the α-curve, and extreme α variations correspond to median β/K values. By increasing the conductivity of the silicon cladding, the amplitude of the curve oscillations decreases slightly, and the α-curve shifts vertically to a higher attenuation. The percent change in attenuation compared to dark conditions ($\sigma = \sigma_0$) for different conductivities ($\sigma = 1.1\ \sigma_0$, $\sigma = 1.25\ \sigma_0$, $\sigma = 1.5\ \sigma_0$) is shown in Figure 4. The β/K curves shift for a conductivity change, as well, although not in such a well-defined manner. The percent phase shift compared to dark conditions ($\sigma = \sigma_0$) for different conductivity changes is shown in Figure 5. These effects will be used for intensity and phase modulation in a device where a signal source induces photoconductivity changes in a thin silicon cladding on a waveguide, thus modulating a coherent beam in the guide.

FIELD CALCULATIONS

We now consider the problem of correlating the local maximum/minimum points on the attenuation and mode index curves with the electric and magnetic field distributions in the waveguide layers. Results indicate that the presence of a thin silicon film (<98Å) has little effect on the wave function profiles; the profiles are similar to those of the three-layer lossless structure (air-dielectric-substrate). For the thick silicon film structure, however, the lowest order mode of the lossless three-layer structure couples to the modes associated with the semiconductor film (the high-refractive index silicon behaves as a waveguide). Furthermore, the coupling between the modes supported by the three-layer lossless structure and the high loss TE' modes* of the silicon waveguide determines the attenuation and phase of the complete four-layer structure.

Our results can be described as periodic coupling between the guided mode and other leaky modes of the same guide. First, we examine the partial structure consisting of a silicon guiding region surrounded by semi-infinite layers of air and polystyrene. The attenuation and mode index are shown in Figures 6 and 7. We note a phase match condition between the modes of the partial structure (air, silicon guide, polystyrene) and the TE_0 mode of the complete waveguide at cutoff thicknesses** for successively higher order modes of the partial structure. The sharp peaks on the attenuation curve, for the four-layer structure, occur whenever the guided wave is strongly coupled into the high-loss modes of the silicon partial structure; conversely, the sharp nulls of apparently zero coupling efficiency occur at thicknesses midway between the values for two adjacent leaky modes of the partial structure. This is similar to the results of power transfer calculations for linearly tapered directional couplers (3, 4). The abrupt transitions on the mode index curve of the complete structure occurs when the phase match condition is satisfied and the guided wave couples into successively higher order modes of the partial structure.

We now consider our results in terms of the electric and magnetic field distributions at the local maximum/minimum points on the attenuation vs. silicon thickness curve. The real part of the TE_0 mode electric field profile in the transverse direction is shown for a cladding thickness, $t_2 = .007$ micrometers (Figure 8(a)). We recall that this thickness is below the cutoff value for the silicon waveguide structure; note that the wave function profile is not appreciably distorted.

For the first local minimum ($t_2 = .05\mu m$), we observe that the field strength at the silicon-dielectric interface approaches zero. We also note the exponentially decaying solutions in the outer, semi-infinite layers as expected. For the first local maximum ($t_2 = .09\mu m$), we observe a sharp peak in the wave function profile at the silicon-dielectric interface as coupling to the TE'_1 mode of the silicon guide occurs (Figures 8(b) and 8(c)).

Similar behavior is noted for the next local minimum/maximum pair, Figures 9(a) and 9(b), ($t_2 = .13\mu m$ and $t_2 = .18\mu m$). The field strength is effectively zero at the silicon-dielectric interface for the local minimum, and a sharp peak in the wave function is evident for the local maximum as coupling to the TE'_2 mode of the silicon guide occurs. We also note that the field begins to oscillate in the semiconductor cladding as we couple into the higher order modes of the silicon waveguide structure.

Again, for the next local minimum/maximum pair ($t_2 = .22\mu m$ and $t_2 = .26\mu m$) the number of field oscillations in the silicon increases. (See figures 9(c) and 9(d).) The increase or reduction in field strength at the silicon-dielectric interface is also apparent. At this local maximum ($t_2 = .26\mu m$), there is the sharp field peak at the interface, but the field decays rapidly through the dielectric indicating almost complete energy transfer. For the other maxima cases considered, the sharp peak was evident at the interface, however, a sizable field was still present in the polystyrene dielectric. This indicates that the local maximum (and likely, minimum) values used for the calculations are not the precise values as in the former case (Figures 9(a) and 9(b)).

---

* TE'_1 denotes guided modes in the semiconductor and TE_1 denotes guided modes in the dielectric.

** Cutoff for the silicon guide occurs when $(\beta/K)_{Si} < n_{polystyrene}$.

The field plots indicate, then, that the attenuation and mode index of the four-layer structure may be explained quite simply as a coupling between the basic three-layer lossless waveguide (air-dielectric-substrate) modes and the high loss TE' modes of the silicon guide. For a local minimum on the attenuation-thickness curve, the field at the semiconductor interface is zero, and for a local maximum, a sizable field is set up at the interface which decays rapidly through the guide and substrate. Finally, the number of field oscillations in the silicon increases as we couple into the higher order modes of the partial structure.

REDUCTION IN ATTENUATION THROUGH USE OF FIVE-LAYER STRUCTURE

Thin dielectric buffer layers have been used to lower the attenuation losses of metal-clad dielectric waveguides (5). These layers are placed between the dielectric core and the metal, and act as buffers to remove a large portion of the field from the metal claddings. We now consider the effect of an SiO_2 buffer layer on the attenuation vs. silicon thickness characteristics.

The result for an SiO_2 buffer layer ($\varepsilon_r = 2.12$) of several different thicknesses is shown in Figure 10. We note the familiar damped sinusoidal behavior and the corresponding reduction in attenuation.

The result for an SiO_2 buffer layer ($t_{SiO_2} = 2000Å$) of several different permittivities is shown in Figure 11. Again we note the damped sinusoidal behavior and the corresponding reduction in attenuation.

Our studies indicate, then, that the attenuation may be reduced significantly with an SiO_2 buffer layer while still preserving the oscillatory behavior of the attenuation curve; more effective reduction is accomplished with a lower permittivity buffer layer. Also, a buffer layer increases the β/K values slightly but decreases the amplitude of the oscillations on the β/K-thickness curve.

CONCLUSIONS

Computer modeling studies on four-layer silicon-clad dielectric waveguides indicate that the attenuation (α) and mode index (β/K) behave as exponentially damped sinusoids as the silicon thickness is increased. The observed effect can be explained quite simply as a periodic coupling between the guided modes of the lossless structure and the lossy modes supported by the high-refractive index silicon. Furthermore, the attenuation and mode index are significantly altered by conductivity changes in the silicon; an amplitude modulator and an intensity modulator have been proposed using these results. Predicted high attenuations in the device may be reduced significantly with a silicon dioxide buffer layer between the semiconductor and the polystyrene guide.

Experimental confirmation of the predicted characteristics is still necessary. A number of thin-silicon film waveguides have been RF sputtered but attenuation measurements to verify the damped oscillatory behavior are forthcoming. Conductivity variations of the silicon should demonstrate the modulation capabilities.

REFERENCES

1. Batchman, T.E.: Investigation of Direct Integrated Optics Modulators - Applicable to Data Preprocessors. UVA/528171/EE80/102 (NASA Grant NSG-1567), Univ. of Virginia, Mar. 1980. (Available as NASA CR-162894)

2. Smith, Robert G., and Mitchell, Gordon C., "Calculation of Complex Propagating Modes in Arbitrary, Plane-Layered Complex Dielectric Structures", E.E. Tech. Report No. 206, University of Washington, National Science Foundation Grant; Eng. 76-09937.

3. Jones, Alan L., "Coupling of Optical Fibers and Scattering in Fibers", J. Opt. Soc. Amer., Vol. 55, No. 3, March 1965, pp. 261-71.

4. Smith, Robert B., "Analytic Solutions for Linearly Tapered Directional Couplers", J. Opt. Soc. Amer., Vol. 66, No. 9, September, 1976, pp. 882-92.

5. Rashleigh, S. C., Planar Metal-Clad Dielectric Optical Waveguides, Ph.D. Thesis, University of Queensland, Brisbane, Australia, May, 1975, pp. 190-220.

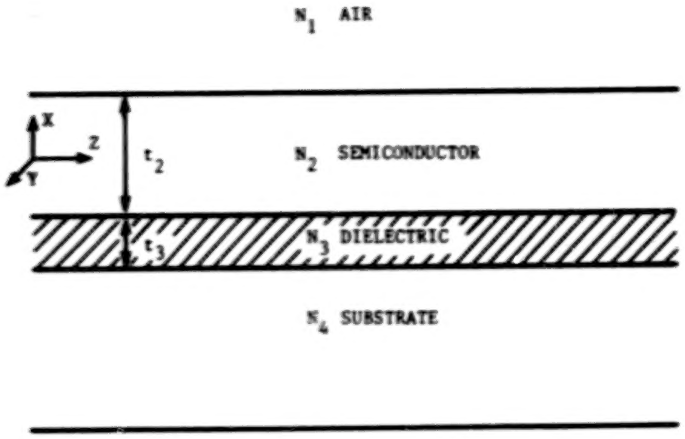

N1 AIR
X
Z
Y
t2
N2 SEMICONDUCTOR
t3
N3 DIELECTRIC
N4 SUBSTRATE

Figure 1.- Four-layer planar waveguide structure.

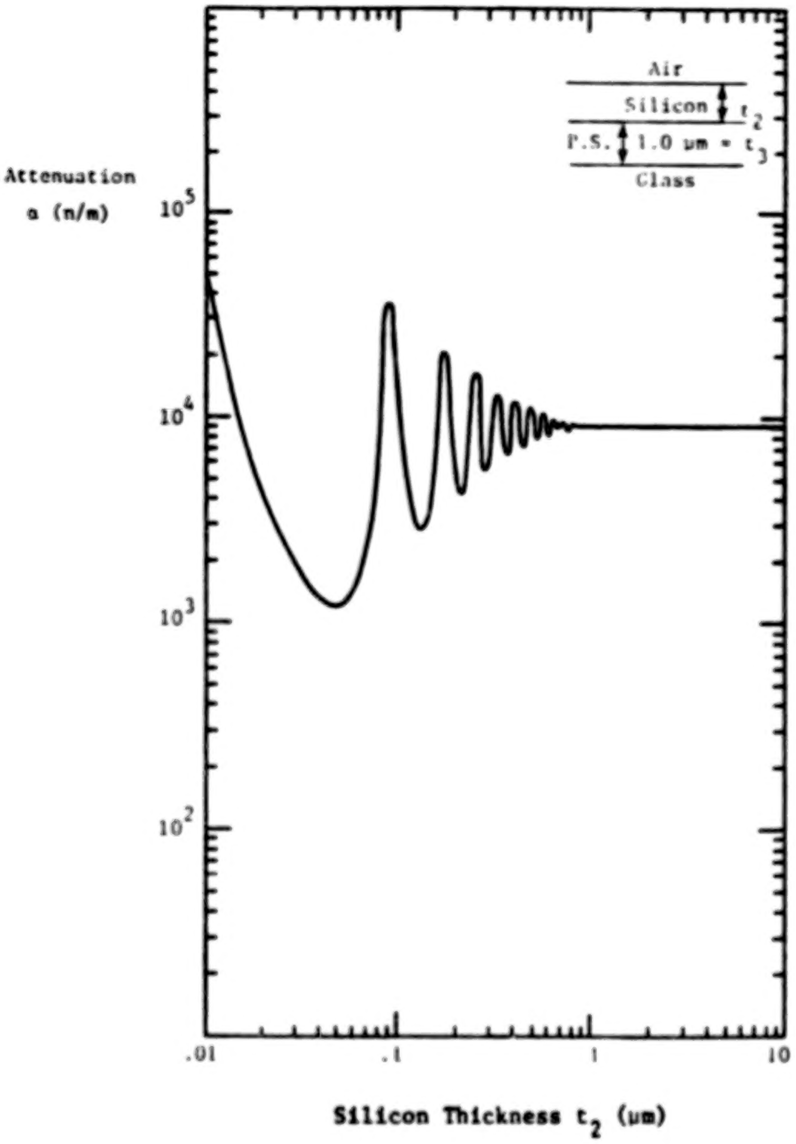


Figure 2.- Attenuation characteristics of silicon-clad waveguide (TE$_0$ mode, normal conductivity).

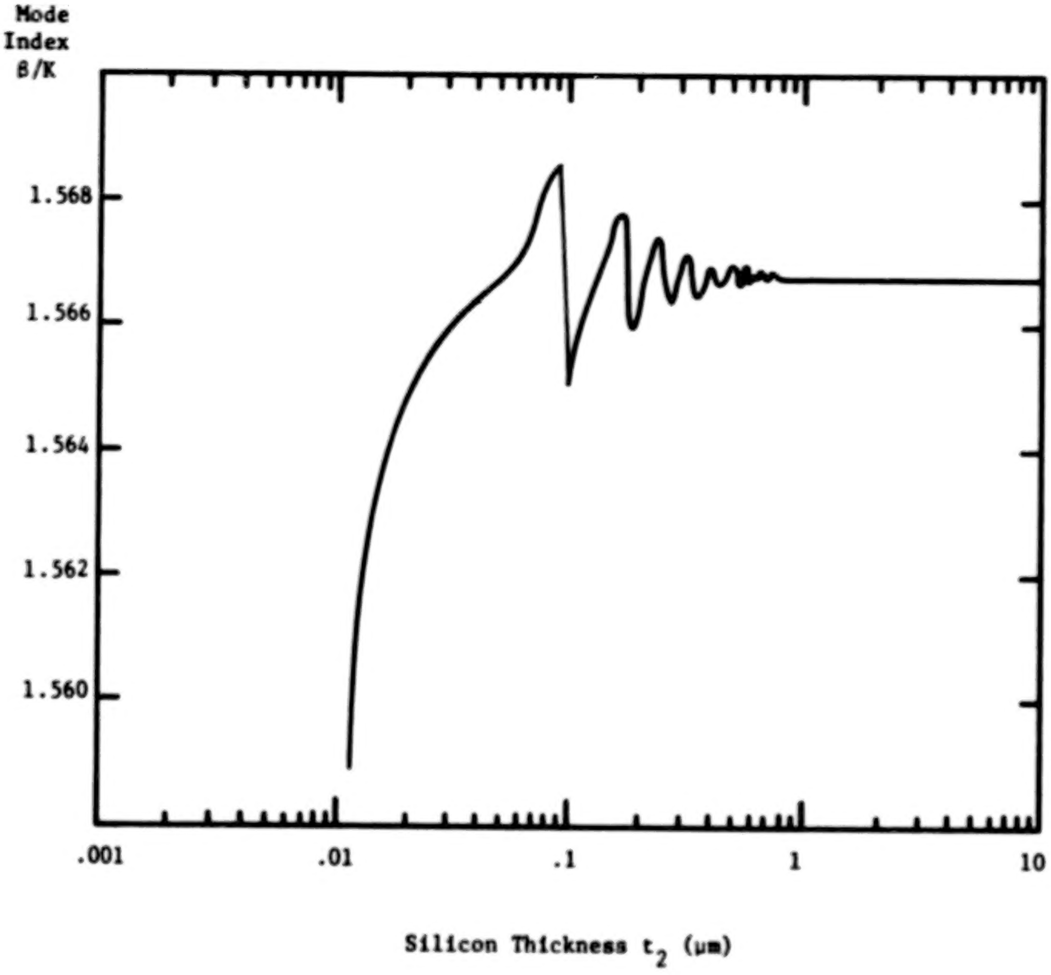


Figure 3.- Mode index characteristics of silicon-clad waveguide (TE_0 mode, normal conductivity).

BEST DOCUMENT AVAILABLE

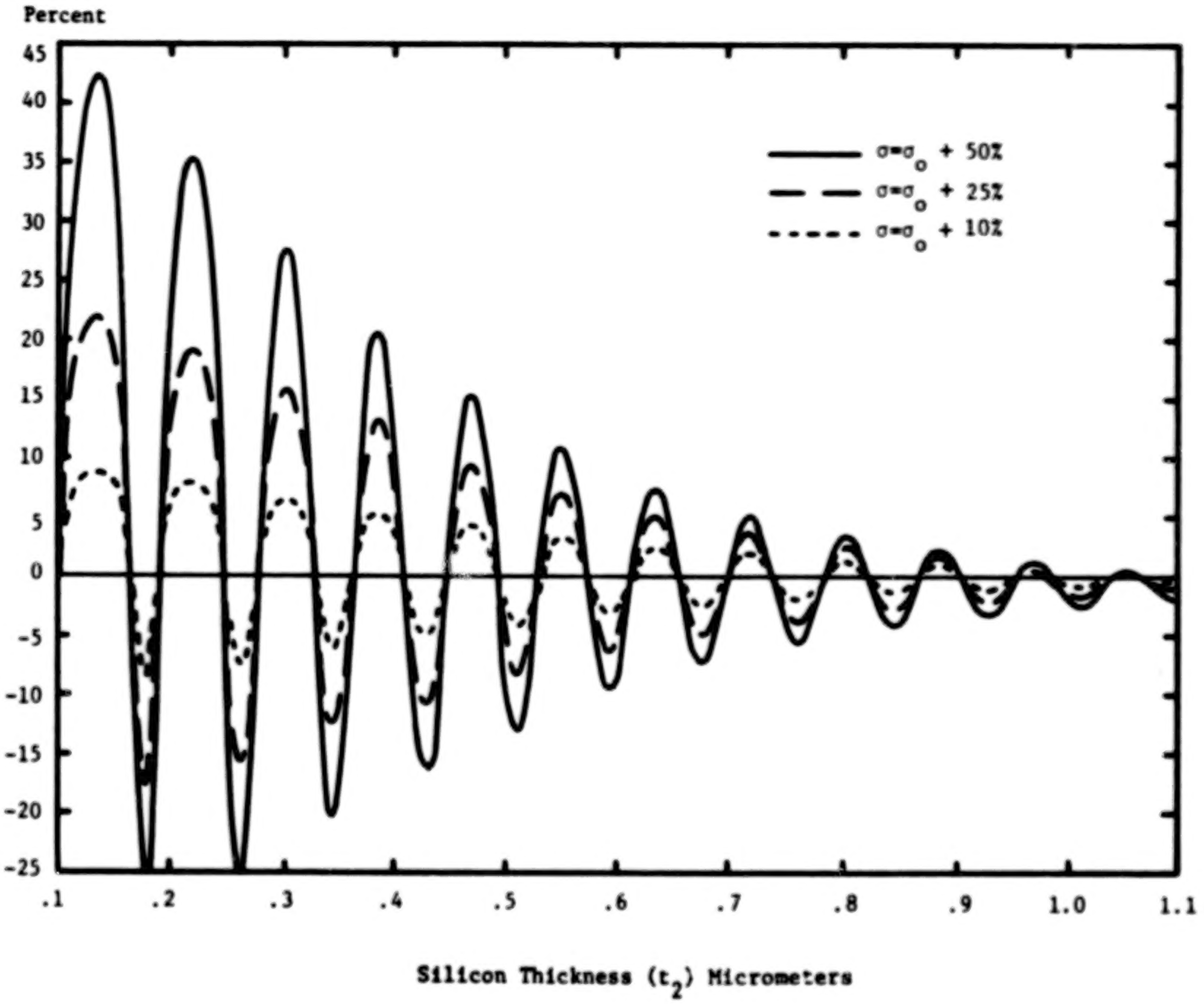


Figure 4.- Change in attenuation with relative change in conductivity (σ_o).

BEST DOCUMENT AVAILABLE

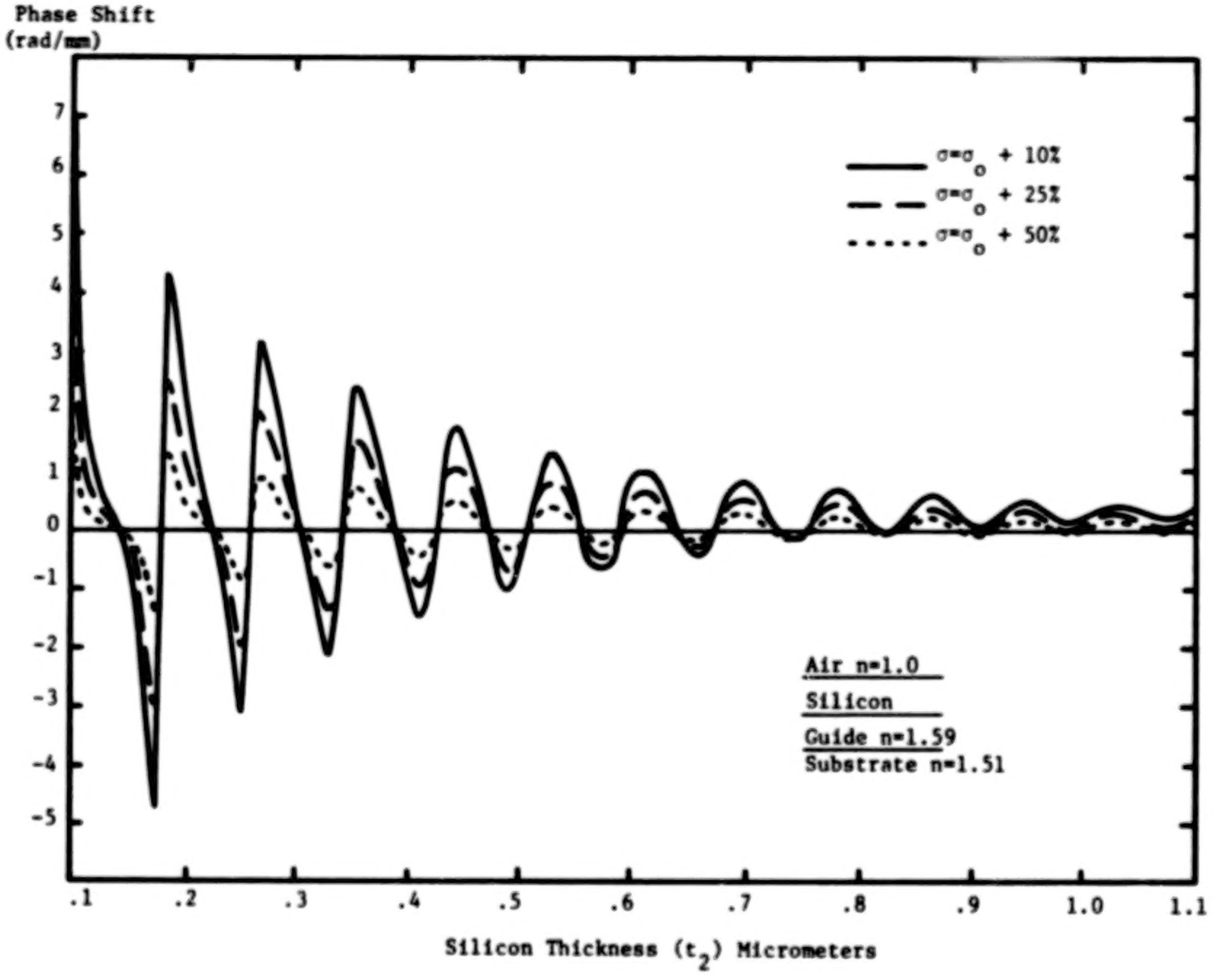


Figure 5.- Change in mode index with relative change in conductivity (σ_o).

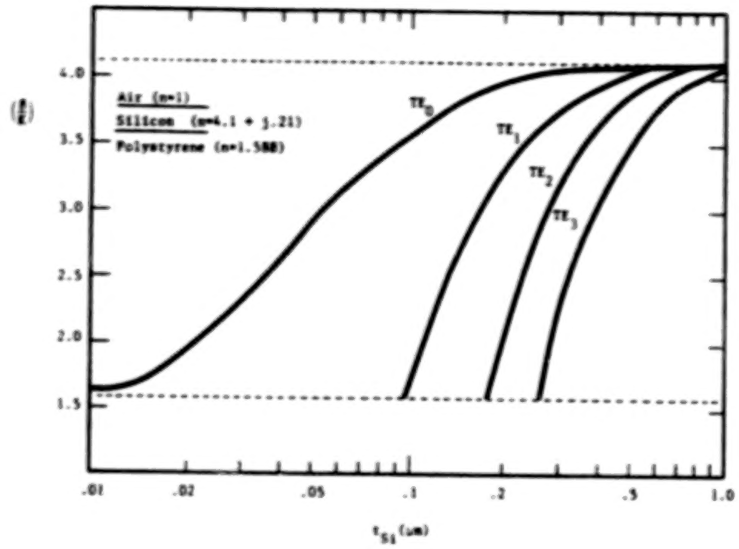


Figure 6.- Attenuation characteristics of silicon waveguide.

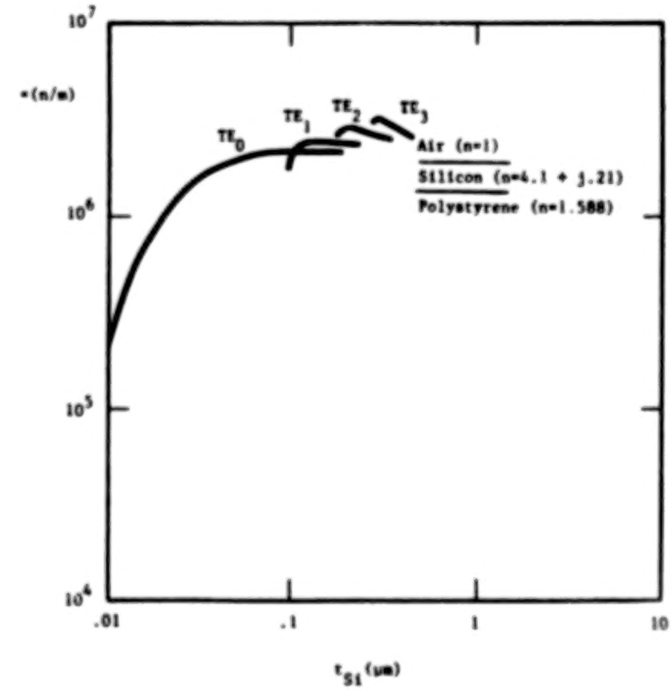
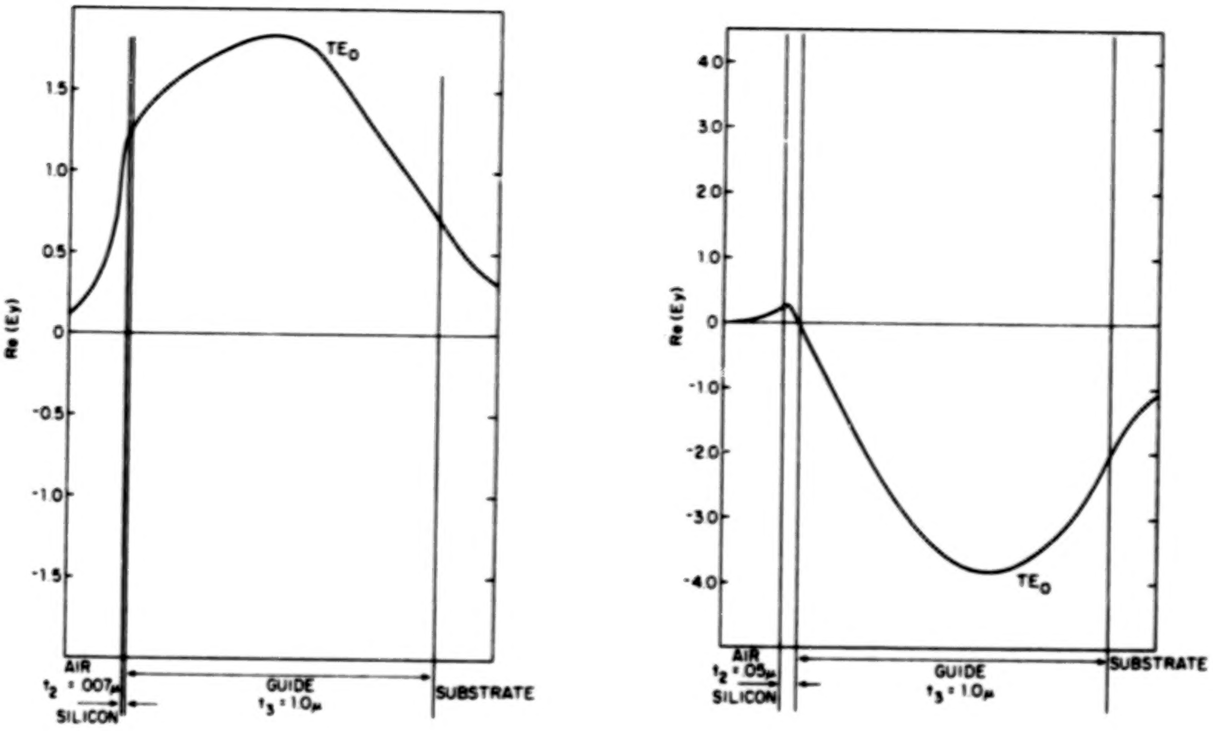
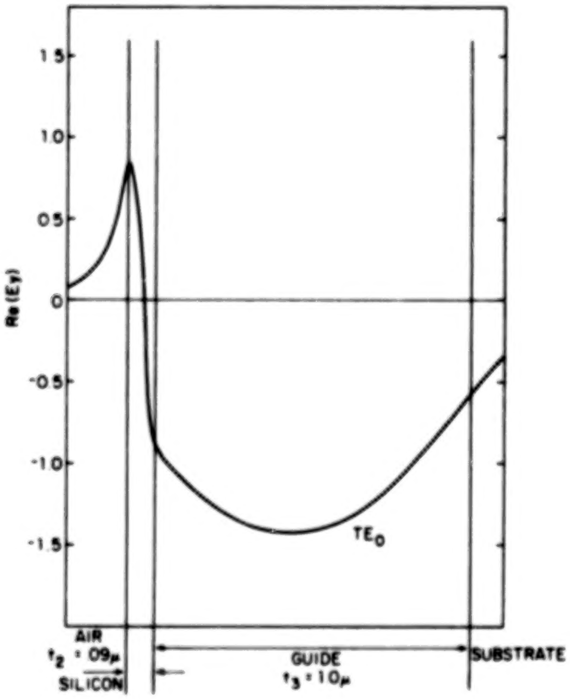


Figure 7.- Mode index characteristics of silicon waveguide.



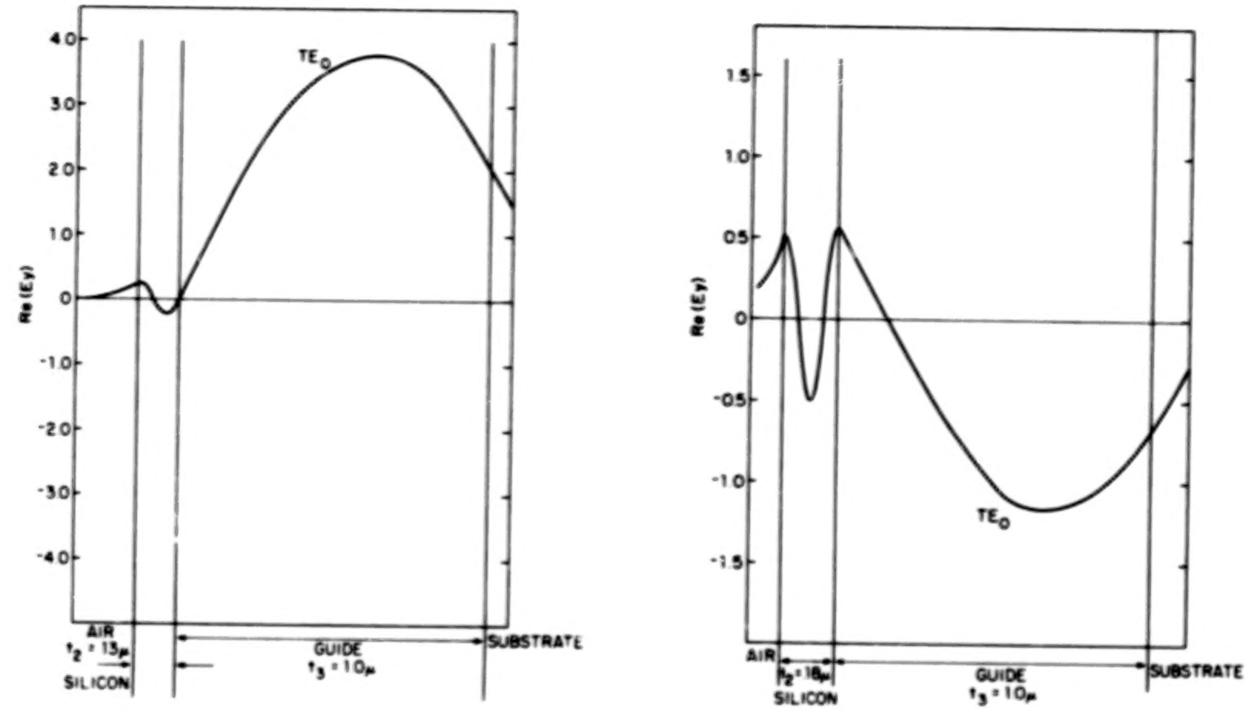
(a) t_{Si} = 0.007μm (below cutoff).

(b) t_{Si} = 0.05μm (local minimum).



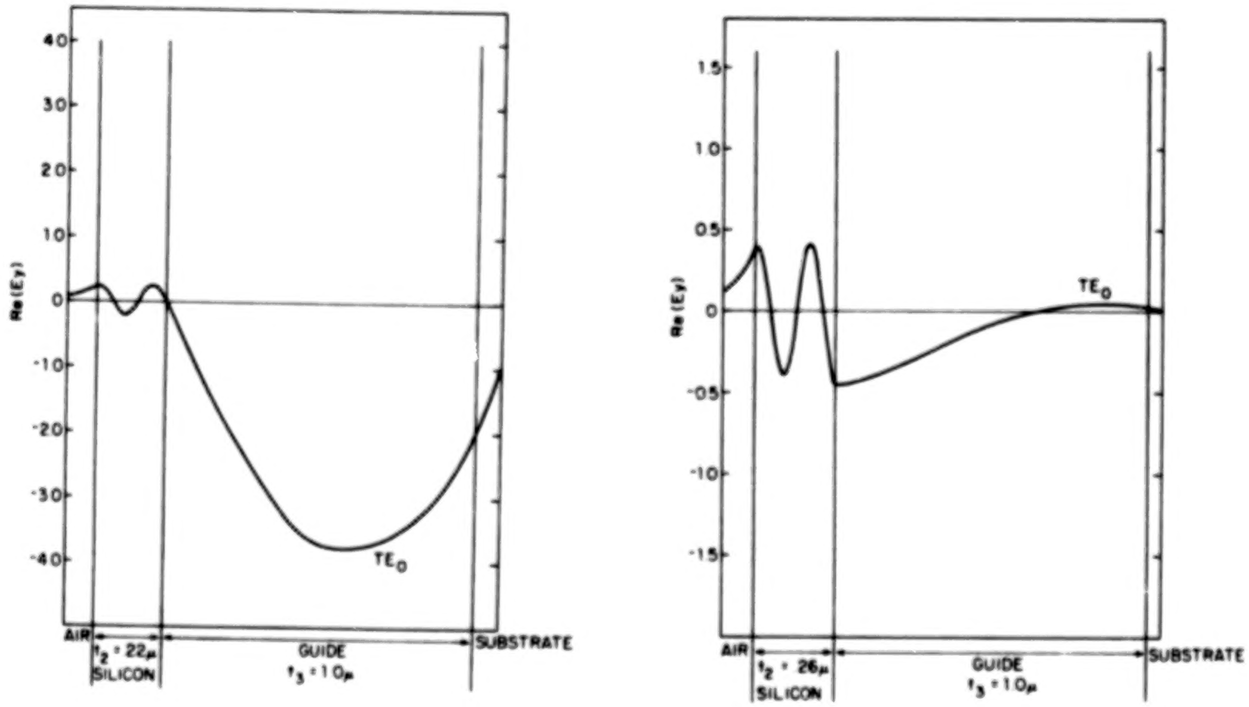
(c) t_{Si} = 0.09μm (local maximum).

Figure 8.- Wave function profile.



(a) t_{Si} = 0.13μm (local minimum).

(b) t_{Si} = 0.18μm (local maximum).



(c) t_{Si} = 0.22μm (local minimum).

(d) t_{Si} = 0.26μm (local maximum).

Figure 9.- Wave function profile.

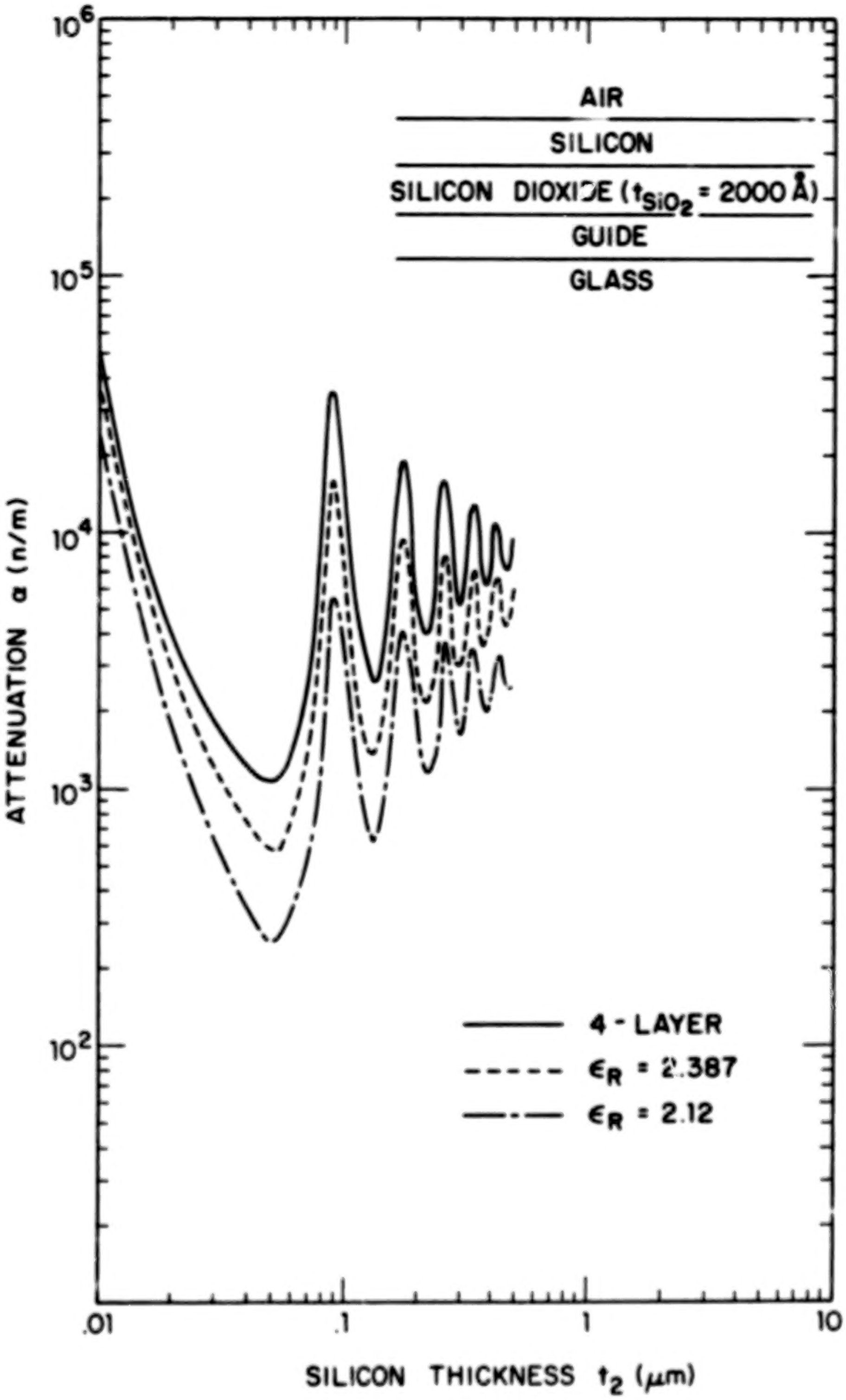


Figure 10.- SiO_2 buffer layer (ϵ_r = 2.12) of different thicknesses.

BEST DOCUMENT AVAILABLE

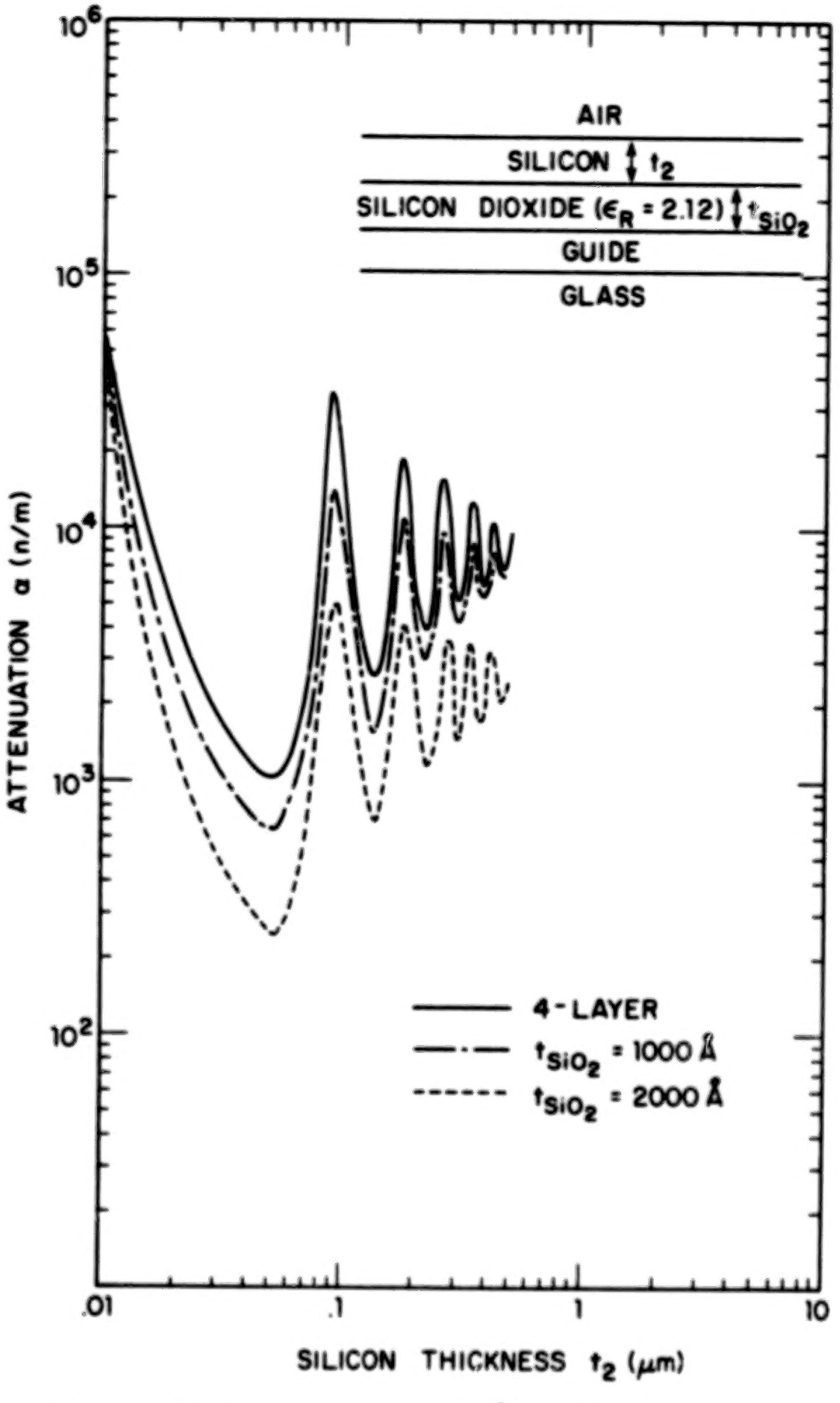


Figure 11.- SiO_2 buffer layer (t_{SiO_2} = 2000Å) of different permittivities.

CO_2 LASER WAVEGUIDING IN PROTON IMPLANTED GaAs

H. A. Jenkinson
US Army ARRADCOM, Dover, NJ 07801

D. C. Larson
Drexel University, Philadelphia, PA 19104

SUMMARY

Surface layers capable of supporting optical modes at 10.6 microns have been produced in n-type GaAs wafers through 300 keV proton implantation. The dominant mechanism for this effect appears to be free carrier compensation. Characterization of the implanted layers by analysis of infrared reflectivity spectra and synchronous coupling at 10.6 microns produce results in good agreement with elementary models.

INTRODUCTION

Integrated optical circuits have vast potential for use as sensing and signal processing elements in future Army fire control systems. One of the requirements for these systems is 24 hour/all weather operation. In order to meet this requirement, the Army has been emphasizing systems which operate at longer wavelengths, in particular, at the 10.6 micron CO_2 laser line. Such systems, including CO_2 laser radars, miss-distance sensors, and sensors for smart munitions, form a natural area of application for infrared integrated optical circuits. These circuits could be used to perform various functions including phased array transmission, heterodyne detection, parallel processing and optical correlation.

Methods which have been used to fabricate optical waveguides and components include thin film deposition, epitaxial growth, ion diffusion, and ion implantation. (ref. 1) The latter method is a particularly interesting technique that can produce electronic, chemical, and optical changes. In crystalline semiconductor materials, ion implantation can alter the refractive index through various physical mechanisms such as damage, doping, stress, and free carrier compensation. However, it is possible, by careful design of the implantation conditions to accentuate one of these mechanisms and to minimize the effects of the others. The present experiments are centered on free carrier compensation of highly doped n-type GaAs through medium energy proton implantation. Although this technique has been used to produce optical waveguides for use in the near infrared (ref. 2), little work has been done to investigate this effect for use at 10.6 microns.

This phenomenon can be illustrated by a simple dielectric model in which the dielectric function of GaAs is composed of Lorentz oscillator terms representing the lattice contribution (ref. 3) and a Drude term representing the depression of the dielectric function due to the presence of free carriers. (ref. 4) In the spectral region between the band gap and the IR phonon resonance, the dispersion due to the lattice contribution is much less than that due to the plasma for heavily doped material. Thus, to good approximation, the dielectric function for these materials may be written, in terms of wavenumber σ, as

$$n^2(\sigma) = K_L - \frac{\sigma_p^2}{\sigma(\sigma + ig)} \qquad 269\ \text{cm}^{-1} \ll \sigma \ll 2\times10^4\ \text{cm}^{-1} \tag{1}$$

where K_L is the high frequency dielectric constant of GaAs, taken to be 10.7. σ_p^2 is related to the carrier concentration N through

$$\sigma_p^2 = \frac{Ne^2}{(2\pi e)^2 m^* \varepsilon_o} \tag{2}$$

where e is the electronic charge, $m^* = .08m$ is the effective mass of electrons in GaAs, and ε_o is the vacuum permittivity. The damping factor, g, leads to losses and is related to the mobility, μ, of the material through

$$g = \frac{e}{2\pi c \mu m^*} \tag{3}$$

where c is the speed of light. During ion implantation, electronic trapping centers are created in the lattice which compensate the free carriers, effectively raising the refractive index. Neglecting losses, the increase in the refractive index is approximately

$$\Delta n \approx \frac{\sigma_{p,s}^2 (1 - N_f/N_s)}{2n\sigma^2} \tag{4}$$

where $\sigma_{p,s}^2$ is σ_p^2 evaluated using the original carrier concentration N_s before implantation, and N_f is the carrier concentration remaining in the layer after implantation.

Polished (100) wafers of silicon doped GaAs were used in these experiments. The free carrier concentration was approximately $4\times10^{18}/\text{cm}^3$. These wafers were implanted at room temperature with 300 keV protons to fluence levels from 10^{13} to $10^{16}/\text{cm}^2$. The implantation was performed at the Naval Research Laboratory. According to the LSS projected range calculations, the compensated region is expected to be sharply defined at a depth of approximately 3 microns. (ref. 5) Such a compensation profile would be unsuitable for optical waveguiding. However, it is experimentally observed that the implantation produces a damage layer with more uniform optical properties extending from near the surface to the projected range.

INFRARED REFLECTIVITY

Prior to waveguiding experiments the optical properties of the implanted samples were characterized by infrared reflectivity measurements. Differential reflectivity measurements, made over the spectral range of 4000 cm^{-1} to 800 cm^{-1}, yielded interference fringes typical of a thin film/substrate structure, as illustrated in figure 1. The interesting features of these curves are the periodicity of the fringes and their increase in amplitude as the measurement progresses to shorter wavenumbers. A first-order model has been developed to aid in deducing the compensation and the thickness of the implanted layer from the reflectivity spectrum. Basically, losses are neglected and the implanted layer is considered to be a film of refractive index n_f sitting atop a substrate whose index has been depressed a small amount Δn by its higher concentration of free carriers. The reference sample is considered identical to the substrate. Expanding the relevant reflectivity equations in a McLaurin series expansion to first order in Δn and using equation (4) yields

$$R_D^2(\sigma) = \frac{R^2_{sample}}{R^2_{reference}} \approx 1 + \frac{4\sigma^2_{p,s}(1-N_f/N_s)}{n_f(n_f^2-1)}\left(\frac{\sin(2\pi t n_f\sigma)}{\sigma}\right)^2 \tag{5}$$

From this analysis, it can be shown that the thickness of the implanted layer can be determined by the separation S of the fringe minima, and the compensation X by the fringe amplitudes:

$$t = (2Sn_f)^{-1} \quad ; \quad X = 1 - N_f/N_s = \frac{n_f(n_f^2-1)\sigma^2}{4\sigma^2_{p,s}}(R^2_{D,max}-1) \tag{6}$$

The carrier concentration of the substrate may also be determined optically by measuring the reflectivity of an unimplanted sample and noting the wavenumber at which the reflectivity dips to a minimum at the edge of the plasma resonance. (ref. 6) At this wavenumber, the dielectric function of the film is very nearly 1. Again neglecting losses, the equation expressing this condition is easily solved for the carrier concentration:

$$N = \sigma_m^2(K_L-1)(2\pi c)^2\epsilon_o m^*/e^2 \tag{7}$$

The more heavily doped the sample is, the more the minimum moves to longer wavenumbers. This technique is sensitive for doping levels on the order of $10^{18}/cm^3$. For example, when $N=2x10^{18}/cm^3$, $\sigma_m=480\ cm^{-1}$;whereas for $N=4x10^{18}/cm^3$, $\sigma_m = 679\ cm^{-1}$. This difference in wavenumber is easily resolved by a spectrophotometer.

The measurements described above were made on a Perkin-Elmer Model 180 double beam infrared spectrophotometer fitted with dual specular reflectance attachments. As shown in table 1, the compensated layers are roughly 3 microns thick. Although the LSS theory predicts about 1 micron of penetration for each 100 keV of accelerator energy, some dose dependence is observed. The low values of the standard deviation, σ_t, are due to the periodicity of the fringes, which in turn indicates the ion-implanted layers can be modeled as dielectric slabs. The determination of the compensation from the reflectivity data is less straight forward. In general, the fringe amplitudes build up at a different rate than can be accounted for by the dielectric model, even in its exact formalism. The values reported represent the average value obtained. The fractional compensation was about 41% for the $10^{13}/cm^2$ implant and began to show saturation with a dose of $10^{14}/cm^2$. Saturation is clearly evident in the samples implanted to 10^{15} and $10^{16}/cm^2$. The standard deviations, σ_x, are larger in this case, due to the larger uncertainties in the reflectivities and the anomaly in the rate of fringe growth. The last two columns of table 1 show the calculated carrier concentration of the films and the refractive indices at 10.6 microns, respectively. The refractive index of the sample implanted to $10^{16}/\text{cm}^2$ is very close to the literature value of 3.275 for undoped GaAs at 10.6 microns. (ref.7)

INFRARED OPTICAL WAVEGUIDING

Since the reflectivity spectra are interpretable in terms of a single dielectric film/substrate structure, the equations for a planar asymmetric dielectric waveguide were used as the waveguide model: (ref. 8)

$$kt\sqrt{n_f^2-(\beta/k)^2} - \Phi_{fo} - \Phi_{fs} = m\pi \tag{8}$$

where $k=2\pi/\lambda_o$, $\beta/k=n_f\sin\theta_f$ and the phase angles are defined by:

$$\tan \Phi_{fi}^{TE} = \left(\frac{(\beta/k)^2-n_i^2}{n_f^2-(\beta/k)^2}\right)^{\frac{1}{2}} \quad ; \quad \tan \Phi_{fi}^{TM} = (n_f/n_i)^2 \tan \Phi_{fi}^{TE} \tag{9}$$

The subscript i refers to either the cover layer o or to the substrate s. Equation (8) is the familiar eigenvalue condition which must be satisfied for optical waveguiding to occur and is expressed in terms of the ratio β/k. This factor can be interpreted as the effective refractive index of the waveguide for each mode and is restricted to values between n_s and n_f.

Infrared optical waveguiding was achieved in these samples using a 1 watt CO_2 laser operating at 10.6 microns. The radiation was focused through an f-10 ZnSe lens on a dove shaped germanium prism which effected coupling into and out of the waveguide. The emerging beam, totally internally reflected from the base of the prism, was imaged as a spot on a thermographic screen. Waveguiding was identified by an absorption notch in the reflected spot, indicating that radiation was removed from the incident beam. Measurements of the angles at which synchronous coupling was achieved were made for both s and p polarizations of the incident beam. From these angles, the effective indices of the guided modes were determined. As two modes were found, it was possible to numerically invert the mode equations to obtain the refractive index and thickness of the implanted layer.

The results of preliminary waveguiding measurements at 10.6 microns are shown in table 2. Due to experimental limitations discussed below, synchronous coupling was only achieved in the two more heavily doped samples, and in those only the 0-order modes were observed although it is believed those samples will also support the TE_1 mode.

The thickness obtained for sample 0519-7 is considerably larger than that obtained by IR reflectivity, although the values obtained for the refractive index compare well. With sample 0519-4 the agreement is better. The thickness obtained from the waveguide measurement is closer to that obtained by IR reflectivity while the values obtained for the refractive index are identical.

Figures 2 and 3 present these results in a different format. These are computer generated plots of the predicted mode structure for a planar asymmetric dielectric waveguide using as inputs for n_f and n_s values obtained from the IR reflectivity spectra. The TE modes are indicated by the solid curves, while the TM modes are represented by the dashed ones.

A horizontal line has been drawn on each figure at the thickness determined by the IR fringe analysis. The effective indices of the modes supported by each sample would be expected to be at the intersection of the mode curves with these lines, as indicated by the Δ symbols on the figures. The corresponding points obtained by the synchronous coupling measurements are shown by the + symbols.

For sample 0519-4 the measured modes are both obtained at larger values of β/k than expected and their separation is smaller, implying they intersect the mode curves at a larger value of thickness. That the waveguide data points remained on the mode curves indicates that the values of the refractive index obtained for the film by both methods are in excellent agreement.

Another point to be made from figure 2 is that regardless of which value is used for the guide thickness, the TE_1 mode should be supported by this sample. The germanium prism used in these experiments had an angle of 54 degrees which limited the lowest value of β/k obtainable to about 2.7. Another prism is being prepared so the entire range of β/k will be accessible. If the TE_1 mode is observed, the extra data point will provide more information about the actual refractive index profile.

The last feature to be noted from this figure is that the mode curves are bounded, in β/k, from n_s to n_f. The refractive index of the substrate for all the samples is 2.44 and is determined by the initial doping level of the GaAs. The refractive index of the film is determined by the carrier concentration remaining after implantation. It must therefore lie somewhere between 2.44 and 3.275, the value for undoped GaAs. Thus as the implant dose is decreased, the mode curves are all squeezed to the left as the upper limit on β/k decreases.

Figure 3 presents the same data for sample 0519-7. The spacing between the TE_o and TM_o modes indicates the layer is considerably thicker than predicted by IR reflectivity. That the mode curves are displaced slightly to the left of the theoretical curves illustrates the lower value of n_f obtained by the synchronous coupling measurements. This sample should also support the TE_1 mode and possibly the TM_1 mode.

DISCUSSION

Several conclusions can be deduced from these results. The first of these is that proton implantation done at the moderate energy of 300 keV produces a damage layer which will support optical waveguiding at 10.6 microns. Although these layers are very thin relative to the wavelength used, the index change produced is so large that the waveguiding condition is easily satisfied. This layer also appears to have very uniform optical properties since empirical results are explainable in terms of elementary models.

There is also very good agreement between the IR reflectivity results and the coupling measurements, indicating the experimentally easier reflectivity analysis has considerable predictive utility. Although it is not fully apparent from the data presented here, it has been found that these are probably complimentary techniques: IR reflectivity appears to be more sensitive to the thickness of the layer than to the refractive index while the coupling measurements yield greater uncertainty in the thickness of the layer and very small standard deviations in the refractive index. It is expected that the refractive index profile of ion-implanted waveguides will be shown to have a transition region between the guiding layer and the substrate rather than a sharp discontinuity as assumed in the models presented here. The complimentarity of the measurement techniques, as well as the systematic differences in the experimental results, may be explained by the different way light experiences the transition region in each of these techniques.

Finally it is noted that the effective indices of the guided modes are strongly dependent on dose, which is easily controlled. This suggests that structures requiring different velocities of propagation, such as delay lines, can be fabricated simultaneously on the same substrate through proper control of the ion beam.

The measurements reported here were made on as-implanted samples. While current work is directed at completing these coupling experiments, the question of losses and post implantation processing to minimize them will be addressed.

REFERENCES

1. Barnoski, M. K.: Introduction to Integrated Optics. Plenum, NY, 1974.

2. Garmire, E., Stoll, H., Yariv, A., and Hunsperger, R. G.: Optical Waveguiding in Proton-Implanted GaAs. Appl. Phys. Lett., vol. 21, no. 3, 1 August 1972, pp. 87-88.

3. Kachare, A. H., Spitzer, W. G., Frederickson, J. E., and Euler, F. K.: Measurements of Layer Thicknesses and Refractive Indices in High-Energy Ion-Implanted GaAs and Gap. J. Appl. Phys., vol. 47, no. 12, Dec. 1976, pp. 5374-5381.

4. Zavada, J. M., Jenkinson, H. A., and Gavanis, T. G.: Optical Properties of Proton Implanted n-type GaAs. SPIE Proceedings, Vol. 276, 1981, pp. 104-108.

5. Gibbons, J. F., Johnson, W. S., and Mylroie, S. W.: Projected Range Statistics. Dowden, Hutchinson & Ross, Inc., Stroudsburg, 1975.

6. Goldsmith, N., and Oshinsky, W.: Vapor-Phase Synthesis and Epitaxial Growth of Gallium Arsenide. RCA Review, vol. 24, Dec. 1963, pp. 546-554.

7. Cheo, P. K., and Wagner, R.: Infrared Electrooptic Waveguides. IEEE J. Quant. Elect., vol. QE-13, no. 4, April 1977, pp. 159-164.

8. Tien, P. K.: Light Waves in Thin Films and Integrated Optics. Applied Optics, vol. 10, no. 11, Nov. 1971, pp. 2395-2413.

TABLE 1.-CHARACTERIZATION OF SAMPLES BY INFRARED REFLECTIVITY
AS-IMPLANTED: 300 keV H^+ @ ROOM TEMPERATURE

SAMPLE	DOSE $1/cm^2$	t μm	σ_t μm	X	σ_X	N_f $1/cm^3$	n_f (@10.6μ)
0519-1	10^{13}	2.64	.06	.41	.10	2.2×10^{18}	2.79
0519-5	10^{14}	3.05	.07	.57	.22	1.8×10^{18}	2.94
0519-7	10^{15}	3.35	.02	.93	.12	2.8×10^{17}	3.22
0519-4	10^{16}	3.31	.01	.98	.19	6.8×10^{16}	3.26
SUBSTRATE						3.8×10^{18}	2.44

TABLE 2.-CHARACTERIZATION OF SAMPLES BY CO_2 LASER WAVEGUIDING
AS-IMPLANTED: 300 keV H^+ @ ROOM TEMPERATURE

SAMPLE	DOSE $1/cm^2$	MODE	β/k	t μm	n_f
0519-7	10^{15}	TE_o	3.071	4.54	3.20
		TM_o	3.026		
0519-4	10^{16}	TE_o	3.087	3.71	3.26
		TM_o	3.014		

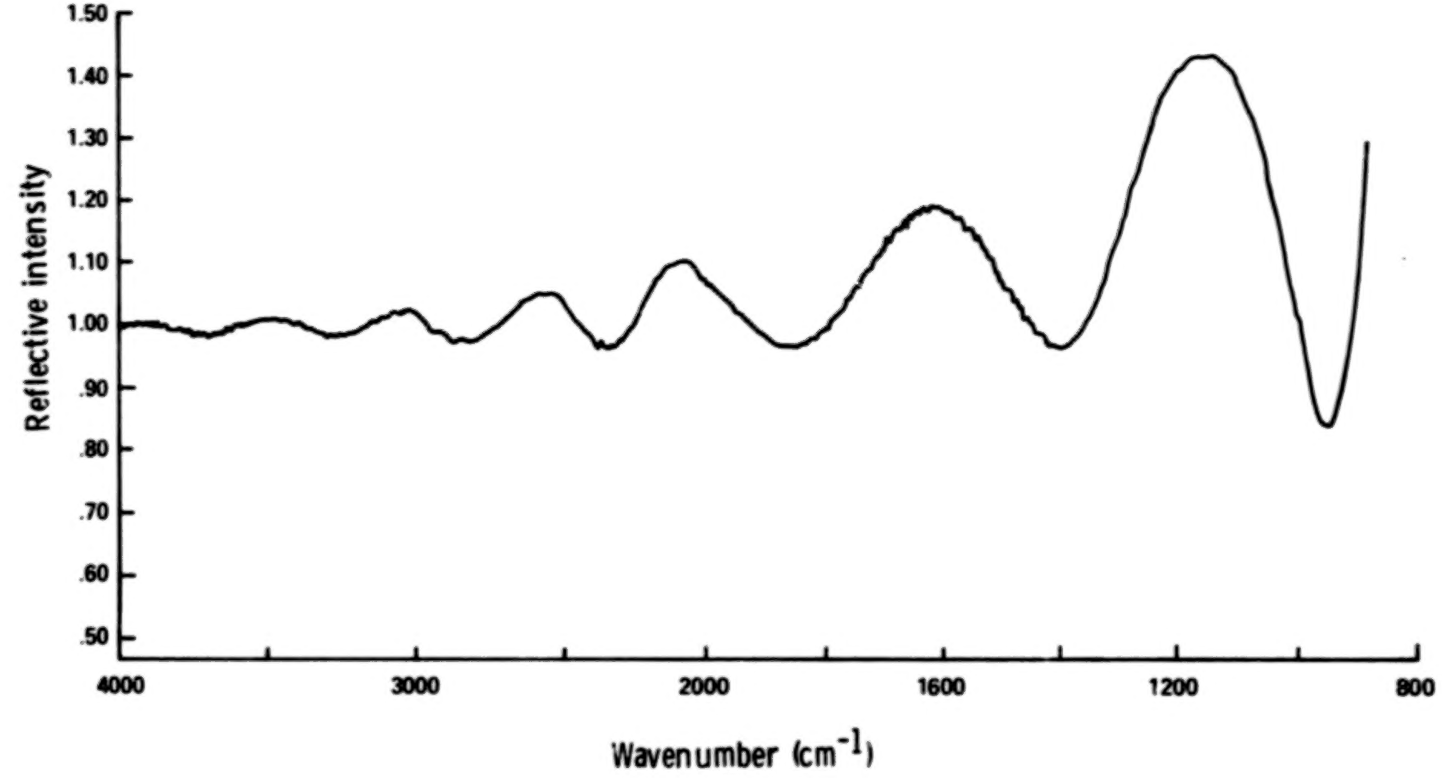


Figure 1.- Differential reflectivity spectrum of sample GaAs 0519-4 (10^{16} H^+/cm^2).

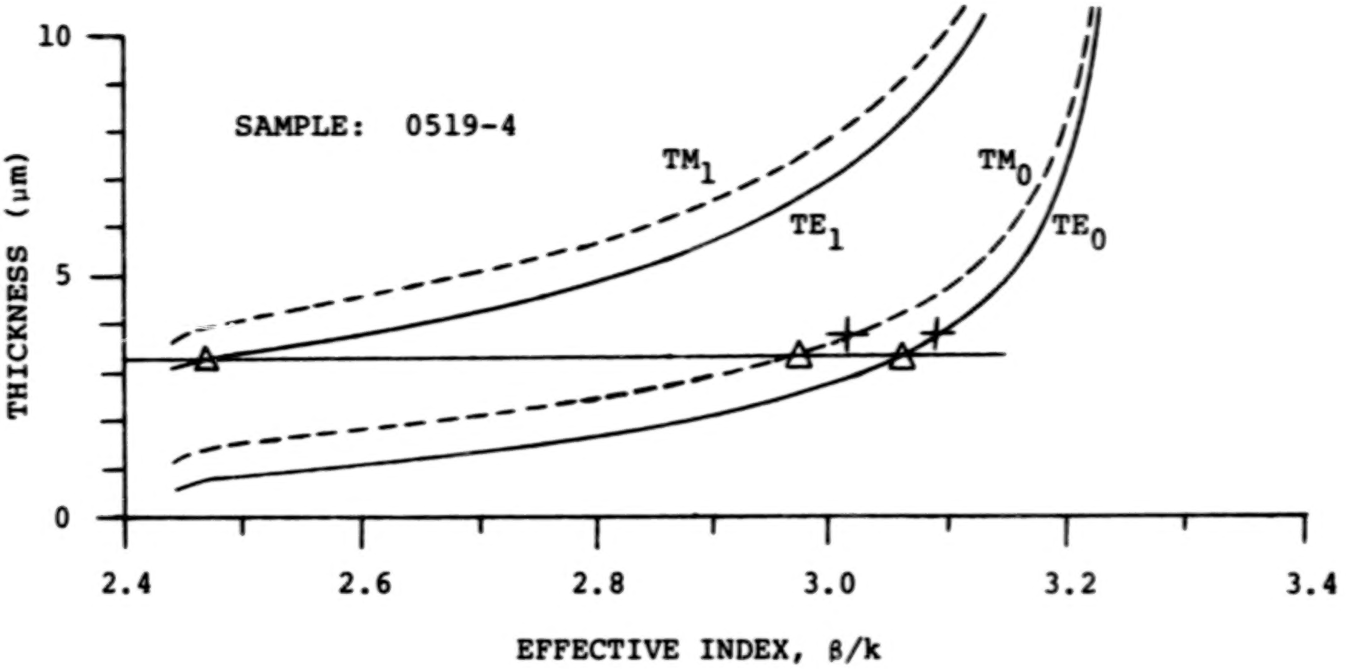


Figure 2.- Comparison of waveguiding and reflectivity results for sample GaAs 0519-4(10^{16} H^+/cm^2). Curves and points indicated by the symbol Δ are predictions based on IR fringe analysis. Results of waveguiding experiments are indicated by the + symbol.

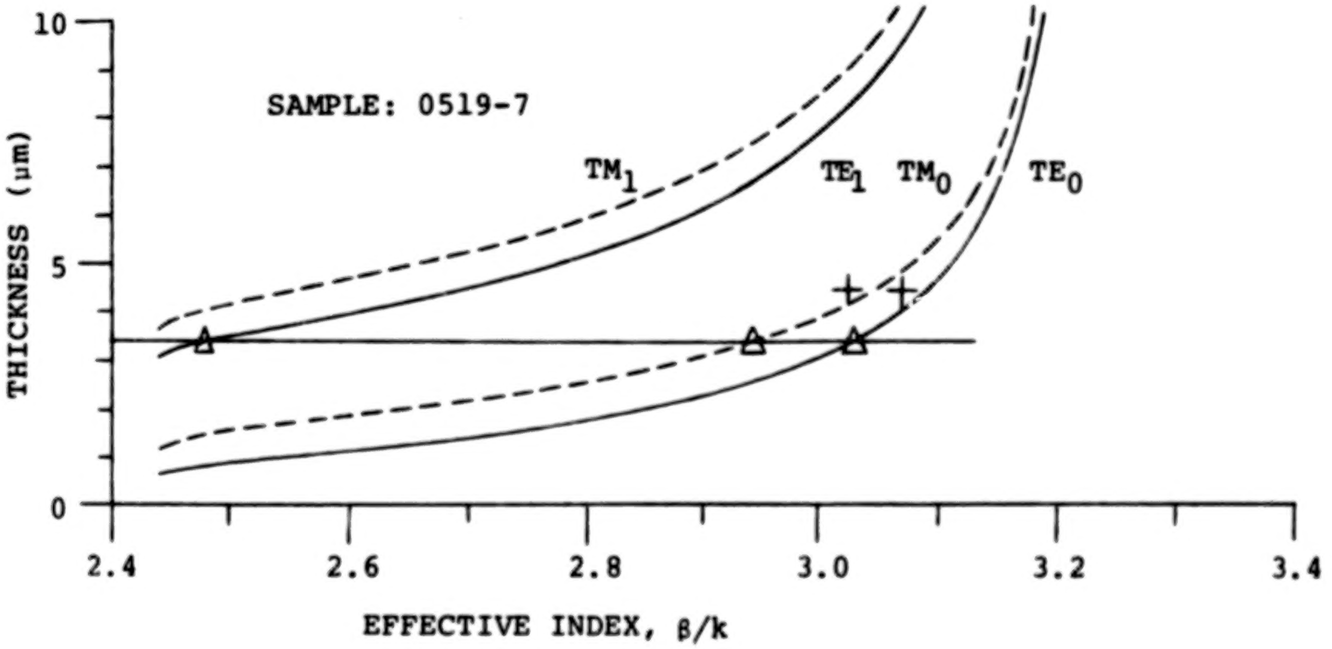


Figure 3.- Comparison of waveguiding and reflectivity results for sample GaAs 0519-7(10^{15} H^+/cm^2). Curves and points indicated by the symbol Δ are predictions based on IR fringe analysis. Results of waveguiding experiments are indicated by the + symbol.

AlGaAs/GaAs TRANSVERSE JUNCTION STRIPE LASERS WITH DISTRIBUTED FEEDBACK

M.J. Hafich, R.A. Skogman, P.E. Petersen
Honeywell Corporate Technology Center
Bloomington, Minnesota
and
H. Kawanishi
Kogakuin University, Japan

ABSTRACT

Two AlGaAs transverse junction stripe lasers, one with distributed feedback and one with a distributed Bragg reflector, have been fabricated. Stabilized single longitudinal mode operation was obtained from both laser structures; temperature dependence of the lasing wavelength was 0.5Å/°K for the TJS/DBR laser and 0.63Å/°K for the TJS/DFB laser.

INTRODUCTION

The transverse junction stripe (TJS) laser in the AlGaAs/GaAs material system combined with a periodic feedback structure such as a distributed Bragg reflector (DBR) has the potential of being a stable, single mode optical source for GaAs integrated opto-electronics and optical communications. Periodic feedback eliminates the longitudinal mode hopping associated with conventional TJS lasers and provides control of the laser oscillation wavelength with very little device to device variation. The TJS laser configuration (see Figure 1.) has both optical and electrical advantages: very low threshold currents (as low as 15mA has been reported for faceted TJS lasers[1]) and single transverse mode optical fields stable over a wide range of currents and temperature.[2,3]

TJS lasers with periodic feedback have been fabricated in two geometries (see Figure 2.)[4]. An interferometric and wet chemical etching technique was used to create a feedback grating across the entire pumping region for the distributed feedback (DFB) TJS laser and to create the separate distributed Bragg reflectors for the TJS/DBR laser. The TJS/DFB laser was a double heterostructure device grown by liquid phase epitaxy (LPE) and had a third order grating etched in the top $Al_{0.2}Ga_{0.8}As$ layer. The grating was buried by growing an $Al_{0.35}Ga_{0.65}As$ layer on the grating by metal organic chemical vapor deposition (MO-CVD). The TJS/DBR laser was also fabricated in an LPE double heterostructure. The top AlGaAs layer was thinned to 0.1μm over more than half of the laser so that the grating would be close to the GaAs active layer and optical field. Single mode operation in both configurations was obtained. The thermal shift of the laser wavelength in both cases was less than 1Å/°K, compared to the 3Å/°K shift of the spontaneous emission peak.

DEVICE PROCESSING

The processing for the TJS/DFB and the TJS/DBR combines conventional AlGaAs/GaAs processing techniques with several special processes. The major steps are shown in Figure 3. For both types of lasers the gratings were in the AlGaAs layer directly above and within 0.1μm of the active layer. This avoids introducing nonradiative recombination of the injected carriers. The $Al_{.20}Ga_{.80}As$ layer of the TJS/DFB laser provides some optical confinement without significantly adding leakage currents around the active GaAs junction. A final layer of $Al_{0.35}Ga_{0.65}As$ was grown by MO-CVD to increase the optical field interaction for optical feedback. The GaAs homojunction was formed by a two step Zn diffusion and ohmic contacts were evaporated on both sides of the laser crystal. Individual devices were cleaved from the wafer with Fabry Perot cavity lengths on the order of 500μm. The inclusion of facets for the Fabry Perot cavity simplifies testing since the structure would oscillate even at a temperature where there was not sufficient gain for the distributed feedback mode to oscillate.

EPITAXIAL GROWTH

The TJS/DBR double heterostructure was grown entirely by LPE. The three layer structure consisted of a 4μm, n-$Al_{0.35}Ga_{0.65}As$ layer; a 0.4μm, n-GaAs active layer; and a 7μm, n-$Al_{0.35}Ga_{0.65}As$ top confining layer. The n-GaAs layer was Te-doped to a carrier concentration of approximately $2\times10^{18}cm^{-3}$. The structure of the TJS/DFB laser was similar except that the top confining layer of the double heterostructure was formed using a hybrid LPE/MO-CVD process. First, an undoped, 0.2μm, n-$Al_{0.20}Ga_{0.80}As$ layer was grown during the LPE process. After the DFB grating was etched into this layer, an n-$Al_{0.35}Ga_{0.65}As$ layer was grown on the grating surface by the MO-CVD technique. Using MO-CVD insured the successful regrowth of AlGaAs that could not be obtained with LPE.

GRATING FABRICATION

The periodic structures for the TJS/DFB and the TJS/DBR lasers were formed by etching corrugations into the top confining layers of the double heterostructure using an interferometrically generated photoresist pattern as the etchant mask (Fig. 4). A HeCd laser (λ=442nm) was used as a light source for the interferometer. Third order gratings with periods between 350 and 400 nm were produced. Very thin photoresist, $\sim$0.1μm, was used to obtain the resolution needed to form the grating. Special care was taken to align the photoresist gratings parallel to the $(01\bar{1})$ direction of the laser crystal. After a photoresist grating had been formed on the AlGaAs surface, a preferential etchant, 3:1:1 H_2SO_4:H_2O_2:H_2O, was used to etch the grating into the confining layer. In the TJS/DBR laser, the pumping region of the laser was protected from the grating etch by a previously applied stripe of photoresist. In both configurations, the process goal was to place the gratings as close as possible to the GaAs active layer without diminishing the gain of the active region by mechanical damage. For the TJS/DFB laser, the photoresist was cleaned after the grating fabrication and a layer of $Al_{0.35}Ga_{0.65}As$ was grown by MO-CVD on the grating surface. The next step in the TJS/DBR process was to sputter Si_3N_4 on the cleaned wafer for a Zn diffusion mask.

ZINC DIFFUSION

The Zn diffusion that forms the p^+-p region of the TJS laser diode was performed in a semi-open ampoule in a two step process. 170nm of low oxygen content silicon nitride was sputtered over the entire wafer and a window for Zn diffusion was opened by Freon plasma etching. The initial p+ diffusion was made at 700°C in a flowing H_2 atmosphere with a $ZnAs_2$ source. Diffusion times up to two hours were required to have the p+ region penetrate to a point just above the GaAs active layer. To form the p-region, the sample was annealed at 900°C in flowing hydrogen with the epilayers lying face down against the quartz boat to minimize material loss. Figure 5 is a scanning electron microphotograph showing the Zn diffusion profile through the double heterostructure. Because the energy bandgap of the lightly doped n-GaAs region was slightly wider than that of the heavily doped p-GaAs active region, the low doped n-GaAs layer in the grating region of the TJS/DBR laser acts as a low loss optical passive waveguide for laser light emitted in the heavily doped p-region.

The final step of the fabrication process was to make ohmic contacts to the n and p regions of the lasers. Evaporated AuGe and Cr/AuZn were used as the n and p ohmic contacts. Both contacts were annealed simultaneously at 460°C. Each wafer was cleaved into separate devices with lengths of 300 to 600 μm. For heat sinking during testing, the lasers were mounted with In solder p-side down on the Au plated copper heat sinks.

RESULTS

A scanning electron micrograph of a finished TJS/DBR laser without ohmic contacts is shown in Figure 6. Figure 7 shows the lasing spectra of a TJS/DBR laser for heat sink temperatures from 133°K to 153°K. In this sample, approximately 9% Al was included in the active layer making the lasing wavelength shorter than that for GaAs. The laser was pumped by applying current pulses with a duration of 200ns and repetition rate of 3Kbits. The laser had a 140μm pumping region with a 220μm long distributed Bragg reflector on one end and a cleaved facet on the other end of the cavity (See Figure 6). Lasing output was measured at the cleaved facet. At 133°K, the wavelength for the DBR mode was longer than the peak wavelength of the spontaneous emission (a few f.p.modes appear) by about 28Å, and the intensity of this mode was not maximum because mismatching occurred between the DBR mode and the spontaneous peak. At 147°K, the matching was obtained and the intensity of the DBR mode increased substantially. The threshold current was 1.3 A at 147°K. Mismatch between the two wavelengths began to appear again as the heat sink temperature was further increased, and the DBR mode wavelength became shorter than that of the spontaneous peak at 153°K. In this temperature range, the DBR mode shifts continuously from 8008Å without longitudinal mode hopping.

The pulsed lasing spectra of a TJS/DFB laser (cavity length of 550μm) for heat sink temperatures varying from 289°K to 307°K is shown in Figure 8. The experimental procedure was similar to that used for the TJS/DBR laser. The thermal behavior of the DFB mode was very similar to that of the DBR mode. The DFB mode was coincidental with the peak of the spontaneous emission at 299°K. The DFB mode oscillated at room temperatures because of an accurate calculation of the equivalent refractive index of the active waveguide and the grating period.

Figure 9 shows the light output vs. injection current (I-L) characteristics as a function of the heat sink temperature for the TJS/DFB laser. At 289°K, where there was a large mismatch between the DFB-mode and the spontaneous emission peak, the I-L characteristic was soft due to the relatively large amount of spontaneous emission. As the temperature was increased to 299°K, the intensity of the spontaneous emission decreased and a much sharper I-L characteristic was obtained. At 307°K, the I-L curve was similar to that of 299°K; however the light output began to saturate with the increase of injection current due to the heating of the active region. This data shows that the differential quantum efficiency was nearly independent of the heat sink temperature.

Temperature changes have two effects on the optical output of TJS/DFB and TJS/DBR lasers. First, the peak of the spontaneous emission shifts with temperature causing the amplitude of the DBR or DFB mode to vary as was shown in the data of Figures 7 and 8. The shift in the spontaneous emission peak was measured to be 3Å/°K which was in close agreement with the shift of the energy gap in GaAs. Secondly, the DFB and DBR modes shift with temperature. The DBR thermal mode shift was measured to be 0.5Å/°K, and the DFB thermal mode shift was measured to be 0.63Å/°K. The stability of the modes of the periodic feedback lasers depends largely on the laser waveguide's refractive index variation with temperature. Calculations of the difference in the waveguide effective refractive index at different temperatures were in close agreement with the experimental data, assuming an index change of $4x10^{-4}$/°C in a simple slab waveguide produces a thermal temperature mode shift of 1Å/°C. Because of the dependence of the refractive index the thermal shifting of the DFB or DBR modes cannot be totally eliminated; however, unlike the mode hopping of the Fabry Perot cavity, the DFB or DBR mode shift was small and continuous.

CONCLUSION

Two new laser structures have been fabricated: an AlGaAs/GaAs TJS/DFB laser and an AlGaAs/GaAs TJS/DBR laser. Both devices operated in a single longitudinal mode; maximum power was obtained at 147°K (threshold current 1.3Å) from the TJS/DBR and at 298°K (threshold current 0.78Å) for the TJS/DFB laser. No mode hopping was observed for the temperature ranges used. The temperature dependence of the lasing wavelengths was 0.5Å/°K and 0.6Å/°K for the TJS/DBR and the TJS/DFB lasers, respectively.

REFERENCES

1. S. Nita, H. Namizaki, S. Takamiya, and W. Susaki, "Single-mode junction-up TJS lasers with estimated lifetime of 10^6 hours", IEEE J. Quantum Electron, vol. QE-15, pp. 1208-1209, Nov. 1979.

2. H. Kumabe, T. Tanaka, H. Namizaki, M. Ishii, and W. Susaki, "High temperature single-mode CW operation with a junction-up TJS laser:, Appl. Phys. Lett. 33, No. 1, pp. 38-39, July 1978.

3. H.Namizaki, Trans. IECE Japan, Vol. E-59, pp. 8-15, 1975.

4. H. Kawanishi, M. Hafich, B. Lenz and P.Petersen, "AlGaAs Transverse Junction Stripe Laser with Distributed Bragg Reflector", Electronics Letters, Vol. 16, No. 19, pp. 738-740, 1980.

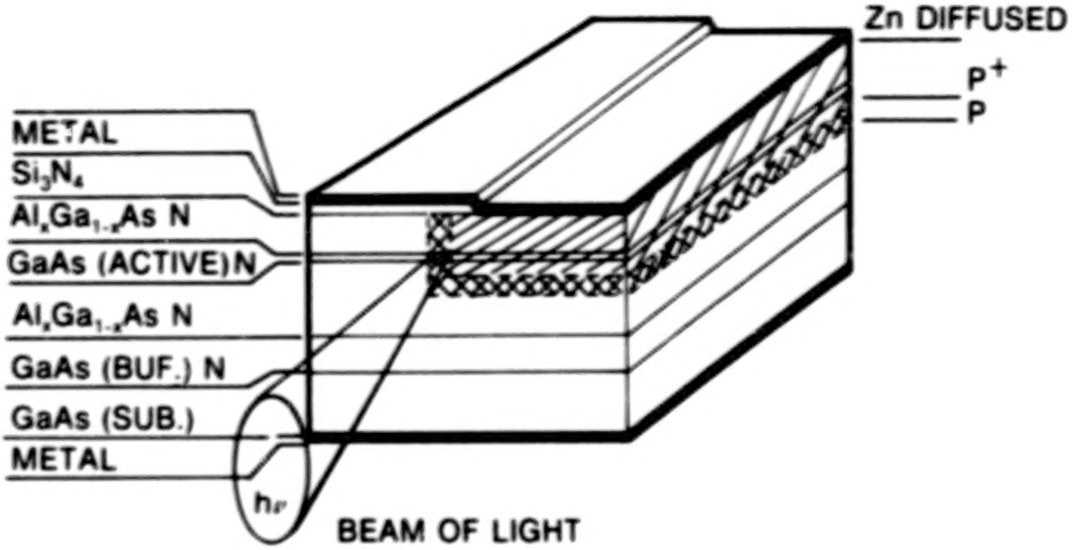

Zn DIFFUSED
P+
P
METAL
Si_3N_4
$Al_xGa_{1-x}As$ N
GaAs (ACTIVE) N
$Al_xGa_{1-x}As$ N
GaAs (BUF.) N
GaAs (SUB.)
METAL
$h\nu$
BEAM OF LIGHT

Figure 1.- Schematic of a TJS laser.

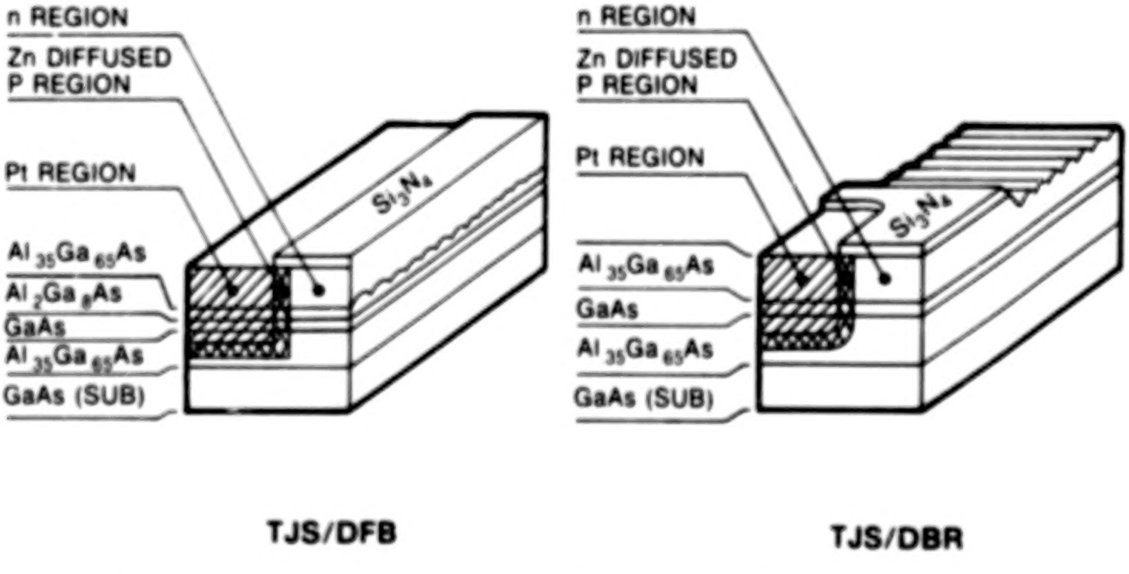
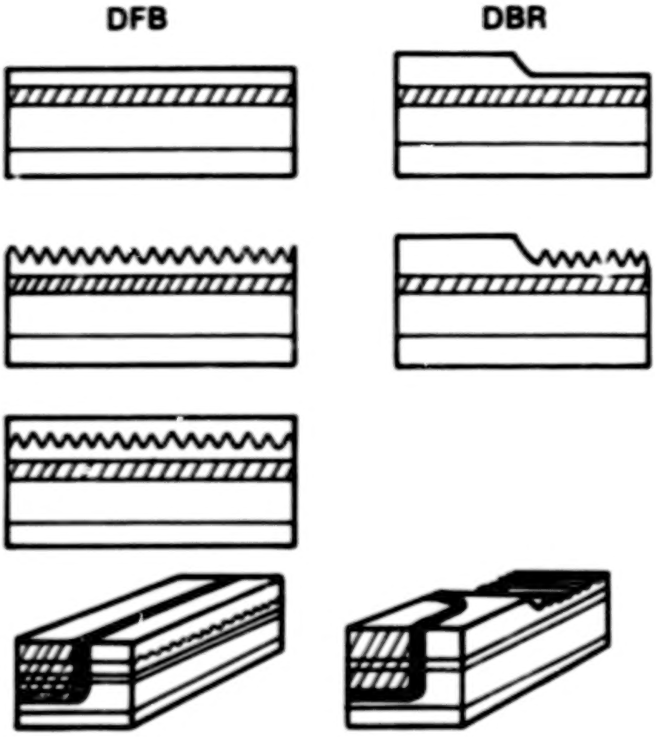

n REGION
Zn DIFFUSED
P REGION
Pt REGION
Si_3N_4
$Al_{.35}Ga_{.65}As$
$Al_{.2}Ga_{.8}As$
GaAs
$Al_{.35}Ga_{.65}As$
GaAs (SUB)
TJS/DFB
n REGION
Zn DIFFUSED
P REGION
Pt REGION
Si_3N_4
$Al_{.35}Ga_{.65}As$
GaAs
$Al_{.35}Ga_{.65}As$
GaAs (SUB)
TJS/DBR

Figure 2.- Structure of the TJS/DFB and TJS/DBR laser diodes.



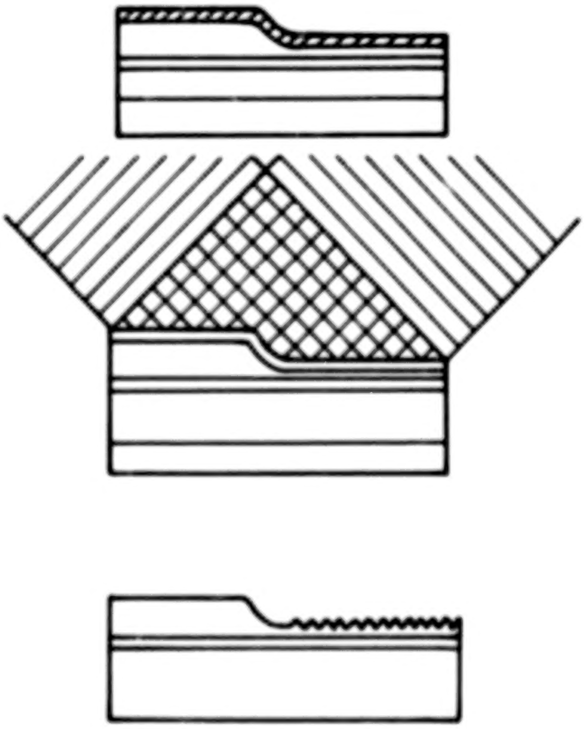
GROW EPILAYERS BY LPE THIN DBR TOP AlGaAs TO WITHIN 0.1 TO 0.2μM OF ACTIVE LAYER.

USE INTERFEROMETRY AND WET CHEMICAL ETCH TO FABRICATE GRATING.

GROW DFB TOP AlGaAs LAYER ON GRATING BY MO-CVD.

FORM P⁺-P REGIONS WITH ZN DIFFUSION. EVAPORATE METAL OHMIC CONTACTS. CLEAVE OUTPUT FACETS.

Figure 3.- Major processing steps for TJS/DFB and TJS/DBR laser diodes.



SPIN COAT LASER WAFER WITH ~ 0.1μM OF PHOTORESIST 1:1, AZ1350B: AZ THINNER

EXPOSE RESIST IN HeCd INTERFEROMETER, λ = 442NM, FOR THIRD ORDER GRATING PERIODS OF 350NM TO 400NM

ETCH GRATING INTO AlGaAs LAYER ALONG (01$\bar{1}$) PLANES. USE 3:1:1, H_2SO_4:H_2O_2:H_2O ETCHANT.

Figure 4.- Grating fabrication process.

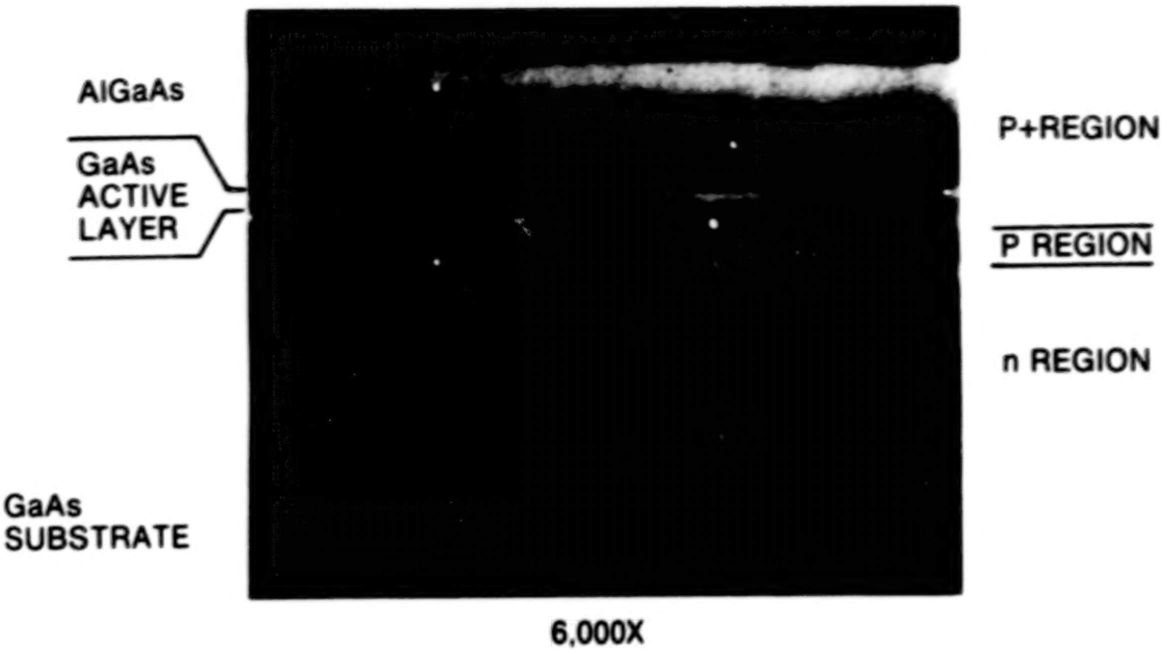

AlGaAs
GaAs ACTIVE LAYER
GaAs SUBSTRATE
P+REGION
P REGION
n REGION
6,000X

Figure 5.- Scanning electron micrograph of the Zn diffusion in an AlGaAs/GaAs double heterostructure.

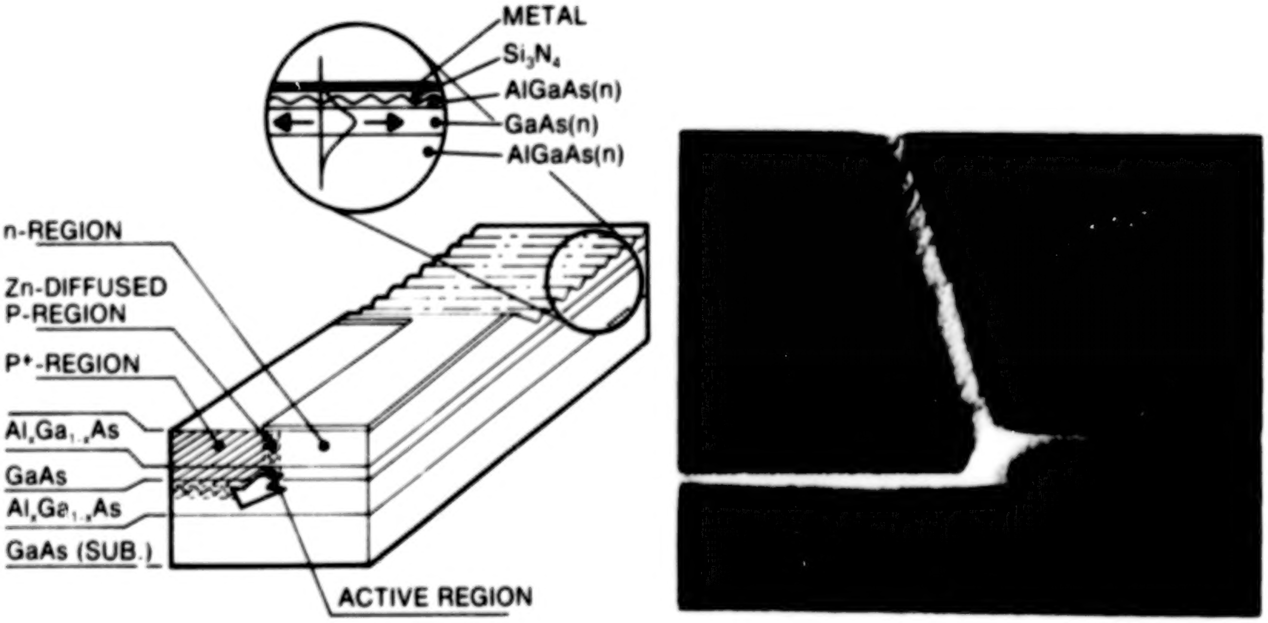

METAL
Si3N4
AlGaAs(n)
GaAs(n)
AlGaAs(n)
n-REGION
Zn-DIFFUSED P-REGION
P+-REGION
AlxGa1-xAs
GaAs
AlxGa1-xAs
GaAs (SUB.)
ACTIVE REGION

Figure 6.- Scanning election micrograph and schematic of the completed TJS/DBR laser.

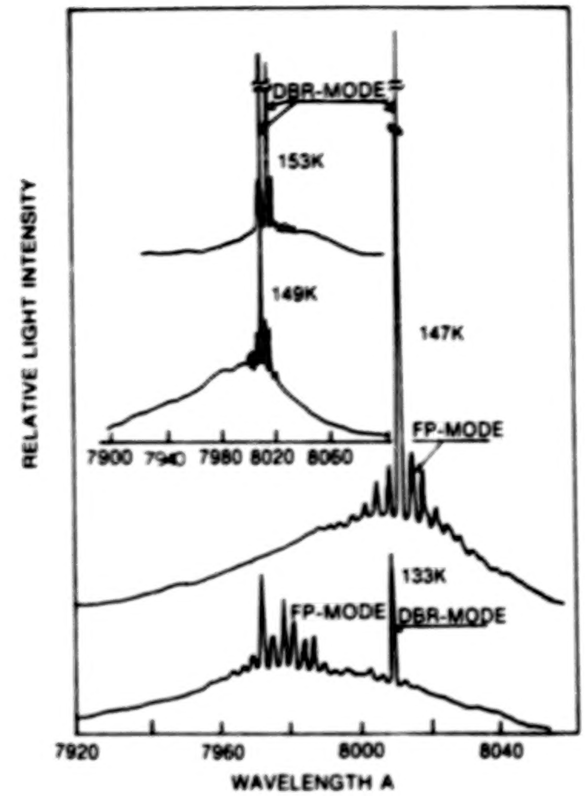


Figure 7.- Mode spectra of the AlGaAs/GaAs TJS/DBR laser between 133°K and 157°K.

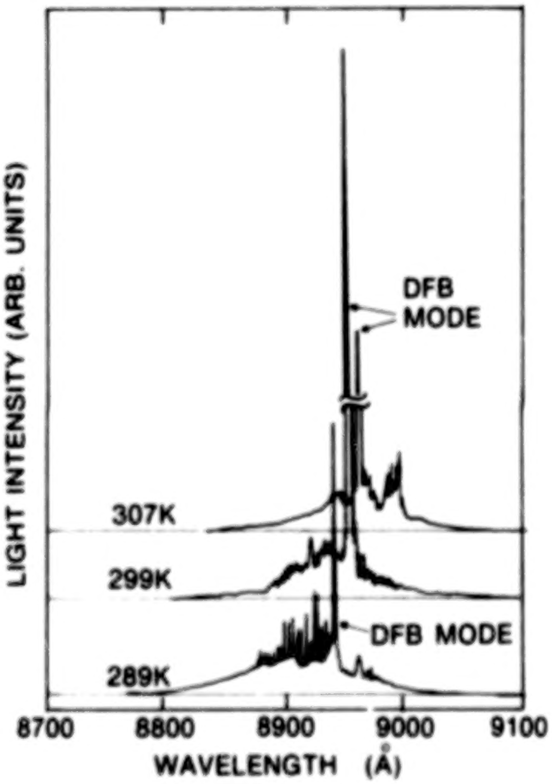


Figure 8.- Mode spectra of the AlGaAs/GaAs TJS/DFB laser near room temperature.

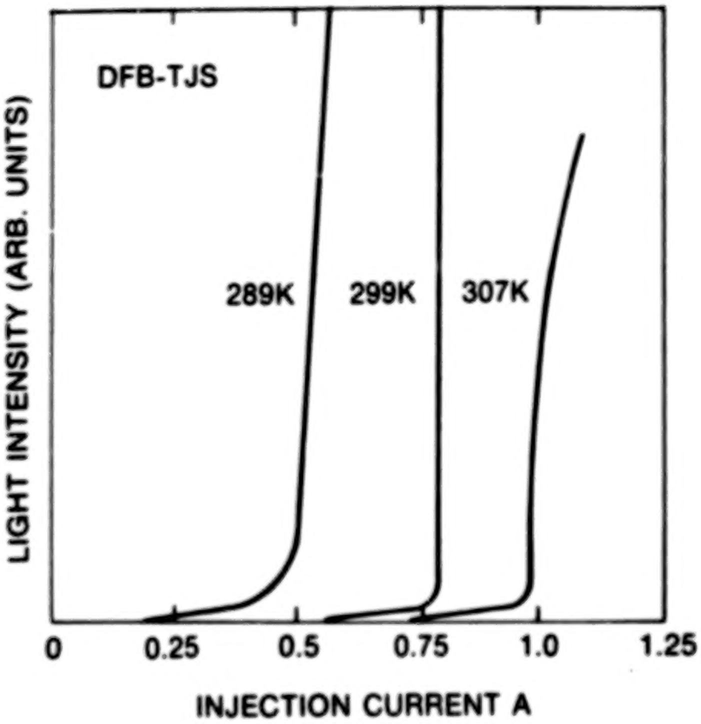


Figure 9.- Light output as a function of injection current (I-L) characteristics of the TJS/DFB laser.

EVAPORATED As_2S_3 LUNEBURG LENSES FOR $LiNbO_3$:Ti OPTICAL WAVEGUIDES*

James R. Busch, Van E. Wood, Richard P. Kenan, and Carl M. Verber
Battelle-Columbus Laboratories
505 King Avenue
Columbus, Ohio 43201

SUMMARY

Because of their relative ease of fabrication, advanced state of design, and low aberrations, Luneburg lenses are attractive candidates as elements in integrated-optics devices, both for simple beam-forming applications and for information processing. Few materials, however, possess the combination of high refractive index, low absorbance, and ease of deposition required for making Luneburg lenses suitable for use on high-index optical waveguides, such as those formed by diffusion of Ti into $LiNbO_3$. We have found that Luneburg lenses of good quality can be formed on such waveguides by evaporation of arsenic trisulfide glass through simple masks. Using only two thin circular-aperture masks, we could readily obtain lenses with focal spots of a few times the diffraction-limited width at f/4. These lenses have been designed for and tested at both visible (633 nm) and infrared wavelengths. Losses are generally small in both wavelength regions, although much of the guided light is coupled into the lens region. There exist restrictions on the steepness of the lens profile near the edge if partial coupling to local modes in the lens region is to be avoided. Since an irreversible change in refractive index can be effected in the As_2S_3 glass by exposure to blue or ultraviolet light, the focal length of the lens can be shortened optically after fabrication. This technique permits in-situ optimization of lens characteristics with no apparent degradation in focal-spot quality. Procedures for the design, fabrication, and testing of lenses of this type will be described.

INTRODUCTION

Realization of integrated optics devices requires development of integrated analogs to the components familiar in conventional optics. Among the most important of such components is the lens. Integrated lenses can be useful not only for the usual focusing and beam-forming operations, but also as data-processing elements exploiting the Fourier-transform property of the lens. We have recently developed techniques for the design, fabrication, and evaluation of one type of integrated lens—the Luneburg lens—on planar optical waveguides formed by diffusion of titanium into $LiNbO_3$. The lenses are formed of overlayers of As_2S_3 glass deposited onto the $LiNbO_3$ surface by thermal evaporation through simple masks. Few materials other than chalcogenide glasses like As_2S_3 possess the high refractive index, low absorbance and ease of deposition required for making this type of lens on a high-index material like $LiNbO_3$. In the present paper we describe and exemplify our current procedures for making As_2S_3 Luneburg lenses. We also describe how the well-known short-wavelength photosensitivity of this material can be used to adjust focal length after lens fabrication.

---

*This work supported by AFOSR and by NASA.

BACKGROUND

Three principal classes of lenses have been investigated for use in integrated optics.

Geodesic waveguide lenses, figure 1a, are formed by creating a spherical depression in the waveguide substrate and then fabricating the waveguide by diffusion or film deposition (ref. 1). A ray of light encountering the depression enters it in a direction determined by the perpendicular distance of the ray from the center of the depression. It then proceeds along a geodesic curve until it encounters the edge of the depression again, whereupon it reenters the planar region of the guide. The index of refraction for the guided mode is practically unchanged in the depression, but the path length depends on the particular geodesic traversed by the ray; focusing occurs because of this differential path length. Since focusing in geodesic lenses is a geometrical effect, all waveguide modes are focused with the same focal length.

Spherical geodesic lenses suffer from spherical aberration which must be corrected for high-performance uses. This can be done by placing a carefully controlled layer of material in the depression, (ref. 2), or by using correcting layers external to the depression (refs. 3,4). Both these methods remove the mode-independence of the focal length. An alternative way to remove spherical aberration is to use an aspheric depression (ref. 5).

Geodesic lenses have several drawbacks. The fabrication techniques are slow and costly. Entrance edges must be rounded to prevent excessive losses and scattering from an abrupt transition from the planar guide into the depression. This rounding tends to lead to an uncertainty in the focal length. Finally, the waveguide in and around the depression must be made stronger because the curvature of the depression tends to introduce radiation losses in weak guides. This is a particular problem in $LiNbO_3$ devices, where crystal size and cost limitations often will necessitate use of short focal lengths.

Diffraction lenses, figure 1b, are waveguide analogs to the Fresnel zone plate. A diffraction lens is a grating whose pattern approximates the thick hologram formed by recording the interference pattern created by the intersection of a converging cylindrical wave and a plane wave. The lens consists of an overlayer of lines that follow the fringes of the interference pattern. In experimental work on such lenses (ref. 6), the line pattern was calculated by computer and replicated by e-beam lithography to generate a mask which was then used to fabricate the lens photolithographically.

The principal advantages of diffraction lenses are low cost, speed and simplicity of fabrication, and potential for automation. Their principal disadvantage is their extremely limited field of view.

The Luneburg lens (ref. 7), is one of the classical gradient-index lenses. It has found application in microwave antennas and beam formers, but until the advent of integrated optics it was impractical in optical systems since, in a bulk optical system, the refractive index at the edge of the lens must equal that of air. A three-dimensional Luneburg lens is conventionally defined as a spherically symmetric refractive index distribution that perfectly focuses the surface of a fixed sphere onto the surface of a second fixed sphere. In the two-dimensional world of integrated optics the distribution is circularly symmetric and the spheres are replaced by circles. Ordinarily we consider the situation where one sphere becomes of infinite radius and the other, in the geometrical-optics approximation, shrinks to a point; so the lens behaves like a conventional lens, bringing a plane-parallel beam to a focus. For specific optical data-processing operations, though, other configurations, possibly not conforming to the conventional definition, might be desired.

The practicality of Luneburg lenses in integrated optics results from the circumstance that the index of refraction at the lens edge is simply that of the waveguide mode rather than that of air. The required refractive-index distribution can be produced in a waveguide by deposition onto the guide surface of an overlayer of material with a prescribed variation in thickness, as indicated in figure 1c. This thickness profile is produced by sputtering or evaporating the lens material onto the waveguide surface through a circular mask with shaped edges. Diffraction-limited Luneburg lenses (ref. 8) have been fabricated using a Ta_2O_5 overlayer on a 7059 sputtered glass waveguide.

Any material that is not highly absorptive can be used to make Luneburg lenses, but only materials with a refractive index higher than the surface index of the waveguide can produce mode-index changes large enough to make short-focal-length lenses. For $LiNbO_3$ waveguides, the surface index is in the range 2.1 to 2.3, depending on wavelength and polarization. Chalcogenide glasses have refractive indices in the range 2.3 to 2.5 and they can readily be deposited by sputtering or evaporation. One of the simplest materials to use is amorphous As_2S_3, which is readily available in the form of chunks of glass from which material yielding good quality films may be evaporated at 500-700°C. We have found this material quite suitable for the formation of Luneburg lenses on $LiNbO_3$.

It is important that the overlay film be tapered in thickness at the lens edges so that spurious modes in the lens region not be excited. These modes occur when the required overlayer thickness is sufficient to allow higher modes in the underlying waveguide to exist. As the film thickness is increased, the single mode of the waveguide is slowly "pulled" up in effective index, until a point is reached when the index rapidly increases beyond the waveguide surface value, toward the bulk index of the film. As this rapid change is occuring, a new mode appears in the waveguide, a mode that was cut off for thinner films. This mode replaces the original single mode of the guide; its index rapidly increases to the vicinity of that of the original guided mode, where it remains until it too is finally pulled up into the film. This process continues, with new modes being pulled up into the guide from cutoff to replace lower modes that have been pulled into the film. As a result, there is, roughly speaking, almost always a mode present having an effective index very near that of the original waveguide. If the film thickness changes too rapidly at the edge, then in the region where the original mode is lifted into the film, strong coupling can occur between the incoming mode and the spurious mode lifted from cutoff by the presence of the film. This coupling occurs because the effective index difference between the incoming mode and the spurious one is small and is changing more slowly than that between the incoming mode and its continuation in the film region. The result of such coupling is the appearance of multiple focused spots and increased apparent loss of transmission. The increased loss may not be too serious, but the transmission of the spurious modes through the lens is deleterious for most applications and can be disastrous for processing applications.

DESIGN

The design of the Luneburg lens involves the following steps:

1. Select the desired operating wavelength, lens aperture and focal length; calculate the f/#.
2. Select the overlay material, waveguide profile, and waveguide parameters.
3. Select the operating mode and polarization.
4. Solve the wave propagation equation for the specified conditions and for an overlay of arbitrary thickness to find mode indices required for propagation both inside and outside the lens region. (As noted by Southwell (ref. 9), it is easier to calculate the overlay thickness as a function of the mode index).
5. For the given f/# evaluate the Luneburg-Morgan integral, which gives the mode index required at a given radius in the lens region to obtain proper focusing.
6. Using the results of 4, determine the lens thickness needed to obtain the desired mode index.
7. From the results of 5 and 6, tabulate the required lens profile.
8. Design a mask arrangement which will produce a lens having a profile close to the specifications of 7.

The first three steps require little discussion. The lenses discussed in this paper are designed for a wavelength, exterior to the guide, of 633 nm. They have apertures of 6 mm and design speeds of f/4.3 for full aperture. Since rays passing very close to the lens edge are not always well focused, we often design for a restricted aperture a little smaller than the lens diameter. The lenses are generally made on single-mode waveguides, for which we assume a Gaussian index profile with a diffusion length of 2 μm. Our design equations are strictly correct only for TE modes; but since for highly asymmetric waveguides the only distinction between TE and TM modes is in the substrate index, good designs for TM modes can also be obtained by simply using the appropriate value for that parameter.

To find the overlay thickness required to obtain a desired mode index (step 4), we have made a straightforward extension of the calculation in reference 9 for a uniform waveguide layer to the case of an inhomogeneous guide characterized by a Gaussian refractive index profile - that is, $n(y) = n_b + (n_s - n_b)\, e^{-y^2/\Lambda^2}$, where n_b is the bulk index of the lithium niobate, n_s is the index at the surface of the waveguide layer, Λ is the diffusion length, and y is depth below the surface into the waveguide. We used the WKB approximation (ref. 10) to determine the mode propagation characteristics of the guide. Since most of the guided light in the lens region propagates in the overlay material rather than in the original waveguide, the exact characteristics of the diffused waveguide do not influence the calculated lens profiles greatly; but this is not to say they do not need to be known with some accuracy. As an example, in one calculation a change of 40% in the assumed surface index n_s led to a change of 10% in the calculated lens thickness at the center.

The required mode-index to obtain focusing at a given lens radius is obtained (step 5) from the solution of the Luneburg-Morgan integral equation.

$$\ln \frac{N(r)}{N_{ext}} = \frac{2}{\pi} \int_0^{(t - z)^{1/2}} \frac{\sin^{-1} (u^2 + z)}{(u^2 + 2z)^{1/2}} \, du. \tag{1}$$

In this equation, $N(r)$ is the mode index at radius r in the lens region, while N_{ext} is the mode index outside the lens. The parameter t is the reciprocal of twice the f/number of the lens, and $z = tR$, where $R = 2r\ N(r)/A\ N_{ext}$. A is the lens aperture. In the form presented here, the integral is easily evaluated to 5 decimal place accuracy by a single 16-point Gaussian quadrature.

Steps 6 and 7 merely involve tabulation and possibly interpolation.

Given the lens profile required for a particular focusing requirement, we need to determine apertures and positions of masks suitable for depositing a lens with this profile on the waveguide. As stated earlier, we have found that fairly good approximations to profiles of interest to us can be obtained using only two thin masks. Our model for the deposition system is a simple generalization of that described in reference 8, and our present design procedure might be termed computer-aided guesswork.

The model for the evaporation system, and the system itself, are illustrated schematically in figure 2.

The evaporation source is modeled as a uniform distribution over a circle of radius b of point sources, each emitting As_2S_3 molecules uniformly into the hemisphere above the source. The extent to which this simple model describes actual evaporation from a crucible has not been investigated experimentally. It is possible that the geometric radius b' of the top of the crucible or the top of the melt differs from the effective source radius b that must be used in the model.

The lens-profiling mask is modeled as a simple circular hole in an infinitesimally thin sheet parallel to the waveguide surface. The distance h_1 of this sheet from the surface and the radius s of the hole are varied by trial and error to optimize the fit to the desired profile. The fits are to a lens profile normalized to unity at the center; we rely on a separate measurement of the central thickness to obtain the proper lens thickness. To obtain a good fit near the edge of the lens, an additional mask very close to the substrate is used. It is positioned experimentally to reduce the lens thickness to zero at the desired aperture.

Presumably with the use of enough additional masks one can approach the design thickness within an arbitrarily small error. More important, though, is the question of how small deviations from the ideal profile affect the operating characteristics of the lengs. We intend to investigate this in future work.

FABRICATION AND CHARACTERIZATION

The As_2S_3 lenses are fabricated in a conventional bell-jar high-vacuum evaporation system. The system is typically pumped down to a pressure of about 1.0×10^{-5} torr by the oil diffusion pump. The source is As_2S_3 fused glass which has been hand ground to form a fine powder. The As_2S_3 powder is thermally evaporated from a quartz crucible held by a conical tungsten basket. The crucible temperature (unmonitored) is estimated to be between 500 and 700°C. The As_2S_3 molecules travel through the series of masks illustrated in figure 2 and are deposited onto the $LiNbO_3$ substrate. Evaporation times vary from 15 seconds to 1 minute and are controlled by a hand-operated shutter. Film thicknesses were estimated during the evaporation process by the weight of the As_2S_3 used as the source and the evaporation time. After the lenses were removed from the system, their thickness profiles were verified by Talysurf. We have recently installed a quartz crystal thickness monitor into the system so that we can accurately monitor the evaporated film thickness in-situ. Care is taken to remove any fumes generated by the evaporation system from the room.

To characterize the optical quality of the Luneburg lenses we have made, we examined the light distribution in the focal plane. The experimental arrangement is shown in figure 3. The input beam was directed through a variable-width aperture and then coupled into the waveguide with a rutile prism. The guided beam passed through the lens and then was coupled out with another rutile prism. The focal plane was external to the waveguide and output coupler. An f/2 imaging lens was used to refocus the focal region onto a diode-array scanner. The scanner consists of 128 photodiodes, each 15 μm wide and 26 μm long, spaced 25 μm apart. The scanner output was directed into an oscilloscope where it was displayed as a plot of detected intensity in the focal plane as a function of position.

EXAMPLES

A lens designed to focus the TM_0 mode of an in-diffused waveguide supporting only one mode of each polarization was fabricated. The source was 0.36 g of As_2S_3 glass in an 18.3 mm diameter quartz crucible positioned 100 mm below the waveguide. The profiling mask had an aperture of 4.7 mm and was placed 22 mm below the waveguide, while the edge-defining mask, with an aperture of 5.9 mm, was 0.5 mm below the guide. A 20 second evaporation yielded a lens with a central thickness of 0.74 μm, close to the design value of 0.75 μm. The measured lens profile as determined by Talysurf is compared with the design profile in figure 4. Although the lens as fabricated was about 20% too thick near the edge, it focused a TM_0 beam only slightly smaller than the full aperture with a focal length of 2.5 cm, close to the design value of 2.58 cm. The flatness of the fabricated lens relative to the design can be alleviated by introducing additional profiling masks. This lens was designed assuming values of 2.293 for the refractive index at the waveguide surface and 2.40 for the As_2S_3 film. More recent measurements indicate that the refractive index of the as-deposited film at 633 nm is at least 2.42 and may be higher. This will not affect the calculated profile greatly, but will have an effect on the design overall thickness.

An example of the focusing qualities of one of these lenses, as determined by the focal-spot scanning procedure previously described, is shown in figure 5. This lens had a central thickness of 0.67 μm and after exposure to an undetermined amount of ultraviolet light had a measured focal length of 2.67 cm. For a one-dimensional lens, the half-power diffraction-limited spot size is given by:

$$S = 0.8859 \frac{\lambda_o}{N_{ext}} \frac{f}{w} \qquad (2)$$

where λ_o is the external wavelength of the light used, N_{ext} is the mode index outside the lens region, f is the lens focal length in the waveguide, and w is the input aperture. We used a 4 mm aperture in this experiment; consequently the minimum half-power spot size is 1.63 µm. The measured value is 5.4 µm; so at this aperture this lens is about three times diffraction-limited. Figure 5 shows that the lens is quite symmetrical with little of the focused beam energy outside the central peak.

The post-fabrication optical shortening of the focal length of the lens by exposure to blue or ultraviolet light is illustrated by the following two examples.

The first lens was fabricated by evaporating from an 81.6 mg charge of As_2S_3 glass for 15 seconds. This 8 mm diameter lens had an original TM_0 focal length of 12.5 cm. After a 60 mJ exposure to ultraviolet light the focal length decreased to 10.5 cm. Subsequent exposures of 30 mJ and 60 mJ shortened the focal length to 9.3 cm and 8.2 cm, respectively. After a total ultraviolet exposure of 540 mJ, the TM_0 focal length had decreased to 5.4 cm.

The second lens was fabricated by evaporating from an 80.6 mg charge of As_2S_3 glass for 20 seconds. This 8 mm diameter lens had an original TM_0 focal length of 5.6 cm. The exposure source was the 4880 Å line of an argon laser. An exposure of 600 mJ shortened the focal length to 4.5 cm. After a total exposure of 2.4 J, the TM_0 focal length had decreased to 3.8 cm, with no apparent aberrations induced by the exposure process.

REFERENCES

1. Anderson, D. B.; Davis, R. L.; Boyd, J. T.; and August, R. R.: Comparison of Optical Waveguide Lens Technologies. IEEE J. Quantum Electron., vol. 13, 1977, pp. 275-282.

2. Spiller, E.; and Harper, J. S.: High Resolution Lenses for Optical Waveguides. Appl. Opt., vol. 13, 1974, pp. 2105-2108.

3. Vahey, D. W.: Corrected Waveguide Geodesic Lenses for Integrated Acoustooptic Spectrum Analysis. Proc. 1978 Ultrasonics Symposium, J. DeKlerk and B. R. McAvoy, eds., (IEEE, New York, 1979), pp. 70-73.

4. Betts, G. E.: and Marx, G. E.: Spherical Aberration Correction and Fabrication Tolerances in Geodesic Lenses. Appl. Opt., vol. 17, 1978, pp. 3969-3974.

5. Vahey, D. W.; and Wood, V. E.: Focal Characteristics of Spheroidal Geodesic Lenses for Integrated Optical Processing. IEEE J. Quantum Electron., vol. 13, 1977, pp. 129-134.

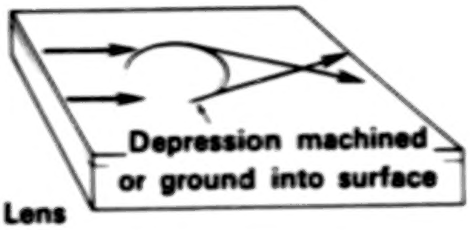
6. Hatakoshi, G.; and Tanaka, S.: Grating Lenses for Integrated Optics. Optics Lett., vol. 2, 1978, pp. 142-144.

7. Luneburg, R. K.: Mathematical Theory of Optics (Brown University, Providence, R.I.), 1944, Sec. 29.

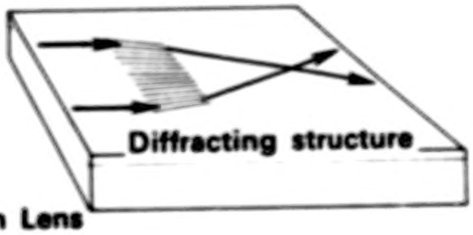
8. Yao, S. K.; and Anderson, D. B.: Shadow Sputtered Diffraction-Limited Waveguide Luneburg Lenses. Appl. Phys. Lett., vol. 33, 1978, pp. 307-309.

9. Southwell, W. H.: Index Profiles for Generalized Luneburg Lenses and Their Use in Planar Optical Waveguides. J. Opt. Soc. Amer., vol. 67, 1977, pp. 1010-1014.

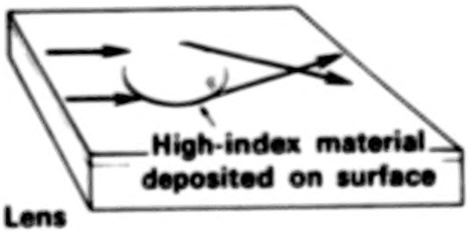
10. Hocker, G. B.; and Burns, W. K.: Modes in Diffused Optical Waveguides of Arbitrary Index Profile. IEEE J. Quantum Electron., vol. 11, 1975, pp. 270-276.


Depression machined
or ground into surface

(a) Geodesic Lens


Diffracting structure

(b) Diffraction Lens


High-index material
deposited on surface

(c) Luneburg Lens

Figure 1.- Three types of waveguide lenses.

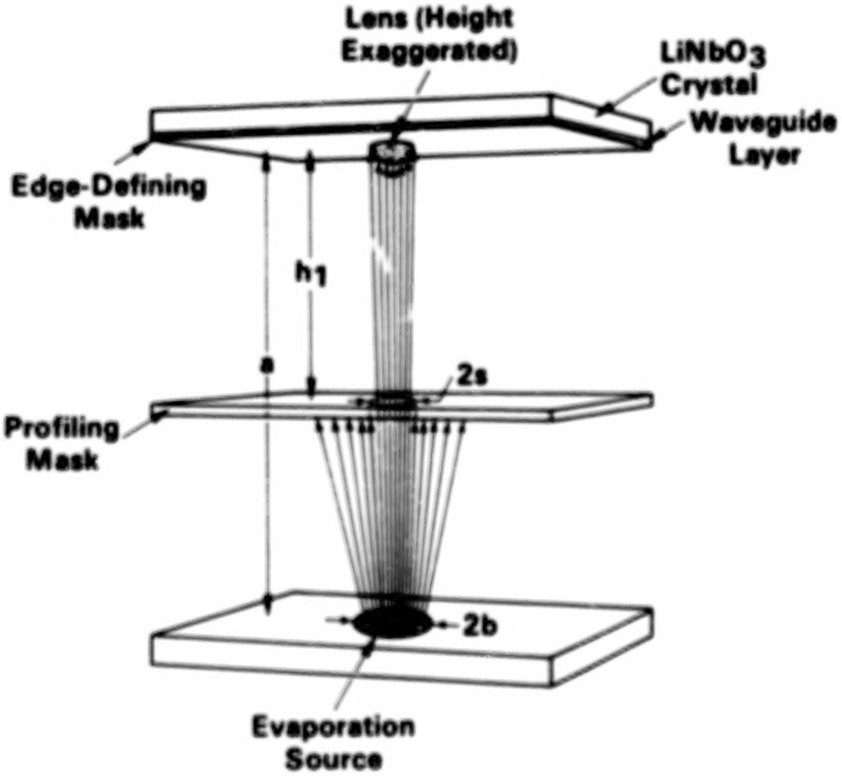

Lens (Height
Exaggerated)
LiNbO3
Crystal
Waveguide
Layer
Edge-Defining
Mask
h1
a
2s
Profiling
Mask
2b
Evaporation
Source

Figure 2.- Model of evaporation and masking system for fabrication of As_2S_3 Luneburg lenses.

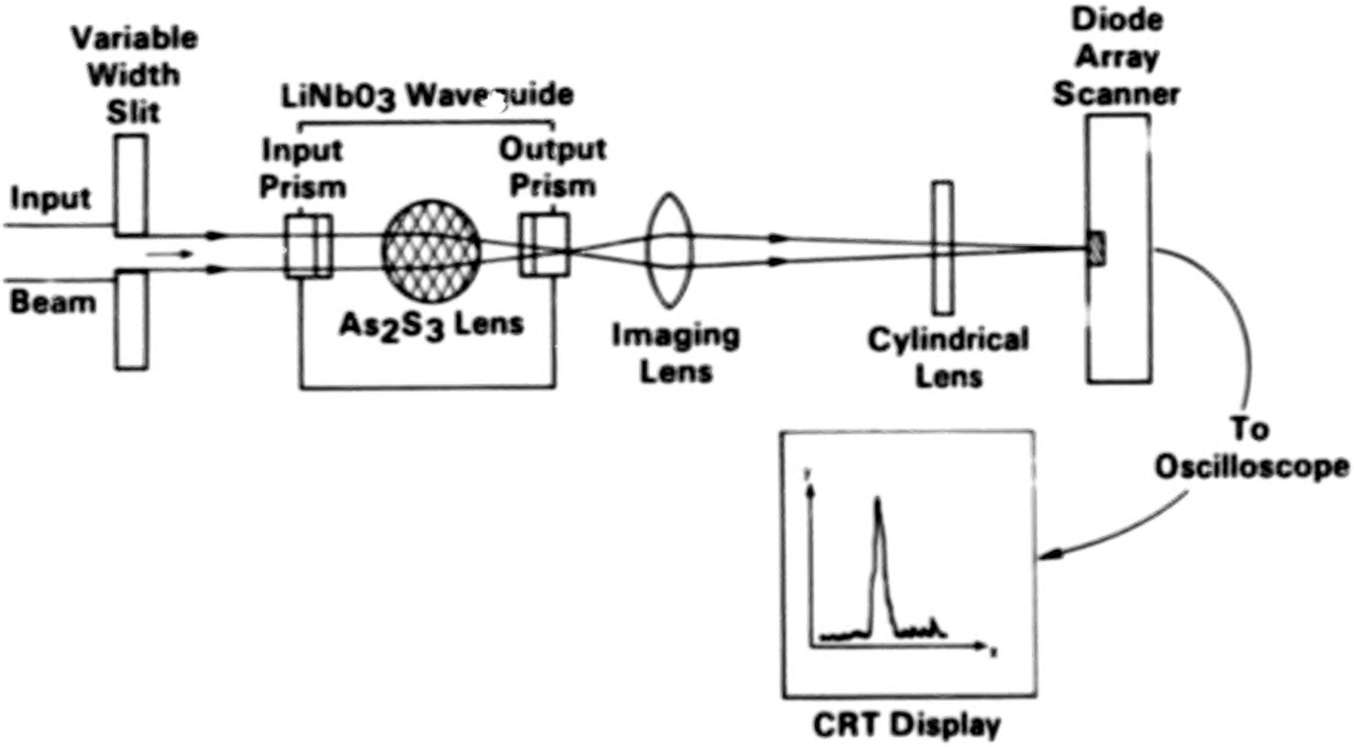

Variable
Width
Slit
Input
Beam
LiNbO3 Waveguide
Input
Prism
Output
Prism
As2S3 Lens
Imaging
Lens
Cylindrical
Lens
Diode
Array
Scanner
To
Oscilloscope
CRT Display

Figure 3.- Experimental arrangement for focal plane scan.

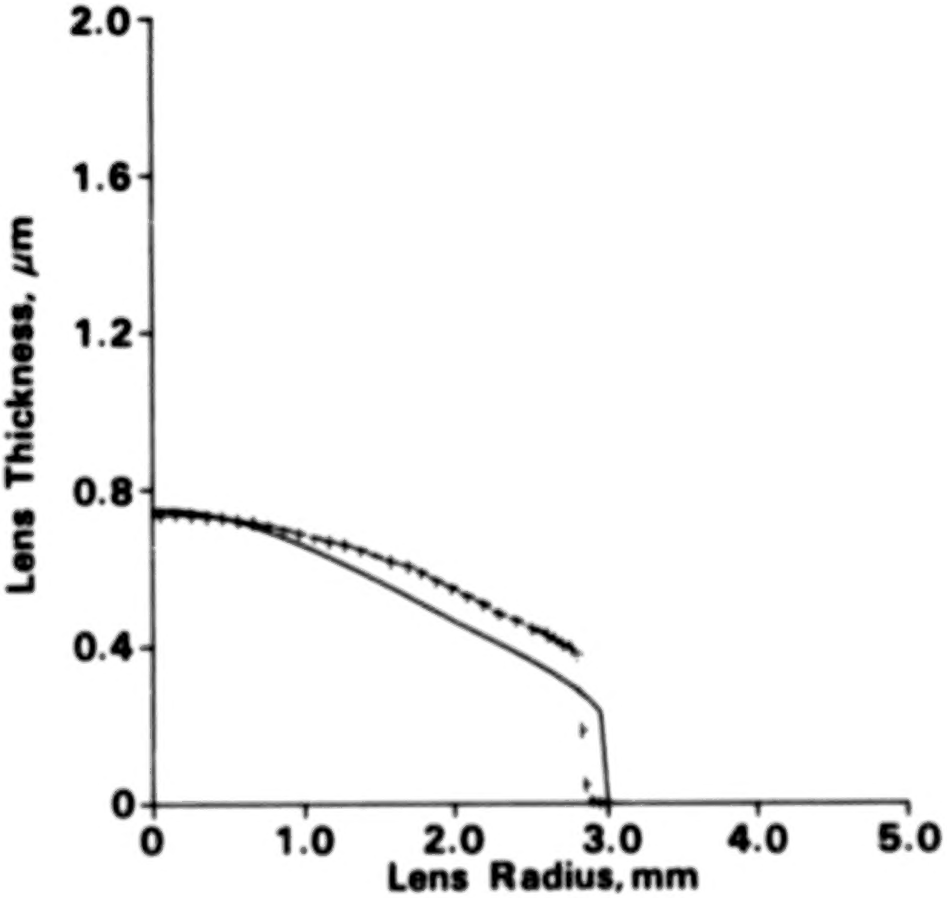


Figure 4.- Comparison of original design profile of Luneburg lens with that of actual lens determined by Talysurf (+'s).

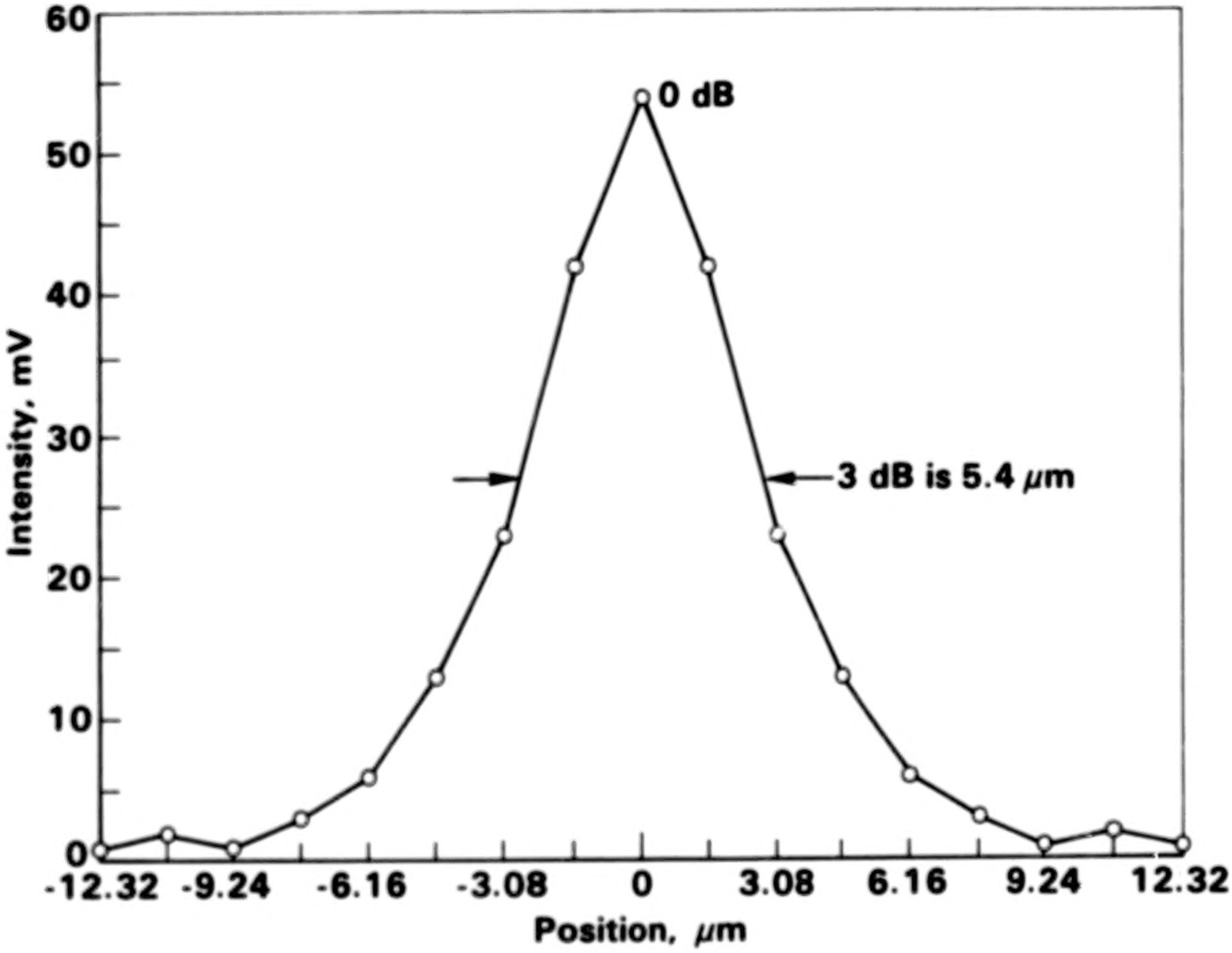


Figure 5.- Energy distribution in focal plane for an As_2S_3 lens of 26.7 mm focal length.

CURRENT AND FUTURE OPTICAL INTEGRATED CIRCUIT TRENDS FOR INFORMATION PROCESSING*

M. K. Barnoski
TRW Technology Research Center
El Segundo, CA 90245

The maturing technology in fiber optics has generated renewed interest in the development of integrated optic devices applicable to single mode guided wave optical systems for data transmission and data processing. Integrated optics represents the development of optical circuits which perform the functions of generation, modulation, switching, coupling, and detection of optical signals in an integrated, compact and rugged format. Many integrated optic devices can be operated at several GHz speed with very low power consumption. The batch fabrication processes developed for microelectronics are readily adaptable to the production of integrated optical circuits, thus reducing the production cost. This presentation will review the status of integrated optic components development and the integration of these devices for information processing. The development of specific optic circuits, such as integrated optic spectrum analyzer, optical correlator, high speed optical analog-to-digital converter, optical switch network, and optical logic circuit, will be reviewed. The presentation will also try to delineate the future trend of integrated optics for data processing.

---

* Because this paper was not available at the time of publication, only the abstract is included.

A BROADBAND RF CONTINUOUSLY VARIABLE TIME DELAY DEVICE

Frederick W. Freyre
Research Laboratories, Hazeltine Corporation

SUMMARY

A method for implementation of continuously variable time delay of broadband RF signals is described. The method uses Bragg Cell and optical heterodyne technology. The signal to be delayed is applied to the Bragg Cell acoustic transducer, and the delay time is the acoustic transit time from this transducer to the incident light beam. By translating the light beam, the delay is varied. Expressions describing the Bragg Cell diffraction, lens fourier transformation, and the optical heterodyne processes are developed. Included are specifications for the variable delay including bandwidth, range of delay, and insertion loss. Applications include radar signal processing, spread spectrum intercept, radar ECM, and adaptive array antenna processing.

INTRODUCTION

Continuously variable time delay of broadband RF signals is a valuable signal processing function. Methods of achieving it include:

- Digital - The signal is digitized, stored, and converted back to analog. This method can yield very long or infinite delays, but is bandwidth limited depending on the number of A/D and D/A converters and their speed. The length of signal delay is limited by the amount of available memory. Digital methods are characterized by large size and power consumption.

- SAW Dispersive Transducers and Reflective Array Compressors - In one version, the signal is applied to a dispersive transducer, then frequency converted, and finally applied to another dispersive element which has equal but opposite slope. The delay depends on the LO frequency used for frequency conversion. Large insertion loss and difficulty in matching the two dispersive slopes are problem areas.

- Magneto-Static Wave Devices - MSW devices exhibit a dispersive time delay characteristic which varies with the applied bias field. Similarly, as with the SAW approach, two dispersive elements having equal but opposite slopes are used. A delay range of 165 to 180 ns and a bandwidth of 250 MHz have been reported (Ref. 1). Difficulties lie in matching the two dispersive slopes.

A new method of achieving broadband variable delay which promises faithful signal reproduction, medium BW, and a wide range of delay uses acousto-optical and optical heterodyne technology. The delay is varied through electronic, not mechanical, means. Insertion loss is inversely proportional to the square of the laser power.

SYSTEM ARCHITECTURE DESCRIPTION

Figure 1 illustrates the system architecture. Included is a laser, a Bragg Cell, a beam translator which can include another Bragg Cell, a Fourier Transform lens, a photomixer and some mirrors and beamsplitters. The broadband signal to be delayed is applied to the Bragg Cell (B1) launch transducer. The resulting acoustic signal then propagates to a laser beam, which is incident at the cell center frequency Bragg Angle. The resulting diffracted light beams are frequency shifted by the acoustic frequencies. A beam combiner combines the diffracted beams and a local oscillator (LO) beam derived from the same laser. A Fourier Transform lens then focuses the combined beam onto a broadband semiconductor photomixer so that optical heterodyning occurs. The result is coherent signal reconstruction for those signal frequency components that are within the Bragg Cell passband and whose diffracted beams interacted with the LO. The time delay is essentially the acoustic transit time from the launch transducer to the incident light beam. Delays ranging from several nanoseconds to several microseconds are feasible. The time delay variation results from translating the incident laser beam along the length of the Bragg Cell while maintaining Bragg angle incidence. Non-mechanical electronic beam translation may be accomplished by using another Bragg Cell placed at the front focal plane of a colliminating lens (Figure 2).

FOURIER TRANSFORM ANALYSIS

At Bragg Cell B1, a family of diffracted light beams is generated since S(t), the signal to be delayed, is broadband. These diffracted beams are frequency shifted by the respective acoustic frequencies. The direction of frequency shift is up if the light is incident on advancing wavefronts, down otherwise. The diffraction angle θ_D for each beam is nearly proportional to the RF frequency. This angle equals $\theta_B + \Delta\theta$ (Figure 3) where

$$\sin \theta_B = \frac{\lambda}{2v_a} f_B \tag{1}$$

and

$$\Delta\theta = \frac{\lambda}{v_a} \Delta f \tag{2}$$

where θ_B = the Bragg angle for the Bragg Cell center frequency and the incident angle
λ = the free space wavelength
v_a = the acoustic velocity
$\Delta f = f - f_B$

The acoustic signal is expressed $S(t + x_1/v_a) \exp - [j\Omega_C(t + x_1/v_a)]$ where v_a is the acoustic velocity and Ω_C the angular RF center frequency. The K vector momentum diagram appropriate to figure 3 (if we ignore beamspread angles) is shown in figure 4. An approximate expression for the amplitude of the diffracted beam outside of the A/O medium is then

$$E_d = \gamma A_d(x_1', y') \, S(t + \frac{x_1}{v_a}) \exp j\omega t \exp - j[\Omega_C(t + \frac{x_1}{v_a}) + \phi] \exp - j(\underset{\sim}{k}_d \cdot \underset{\sim}{r}) \tag{3}$$

where γ relates to the amplitude acousto-optic diffraction efficiency and is assumed constant over the signal bandwidth, $A_d(x_1', y')$ is the diffracted optical beam amplitude distribution in the rotated XZ coordinate system (figure 5), ω is the optical angular frequency, and ϕ is the acoustic signal phase.

In this discussion A_d is assumed Gaussian and is expressed

$$A_d(x_1', y') = \frac{a}{\sqrt{\pi}} \exp - a^2(y'^2 + x_1'^2) \tag{4}$$

where $a = 2/D$ and D is the optical beam diameter. From figure 5 we have

$$x_1' = -Z \sin \theta_d + X \cos \theta_d \tag{5}$$

$$y' = y \tag{6}$$

where θ_d is the Bragg diffraction angle outside the medium. From figure 4 we have

$$\underset{\sim}{k}_d \cdot \underset{\sim}{r} = k(z \cos \theta_i - x_1 \sin \theta_i) + K x_1 \tag{7}$$

where k and K are the free space optical and acoustic wavenumbers, respectively.

The spatial FT of (3) may be expressed:

$$\begin{aligned} FT(E_d) = {} & \gamma \exp j[(\omega - \Omega_C)t - \phi] \exp j\, k\, F \cos \theta_i \frac{a}{\sqrt{\pi}} \int \exp - a^2 y^2 \exp 2\pi v y \, dy \cdot \\ & \int S(t + \frac{x_1}{v_a}) \exp - a^2 (F \sin \theta_d + x_1 \cos \theta_d)^2 \exp - j \frac{\Omega_C x_1}{v_a} \cdot \\ & \exp j(k \sin \theta_i - K)x_1 \exp - j\, 2\pi\, u\, x_1\, dx_1 \end{aligned} \tag{8}$$

where $z = -F$
F = the lens focal length
$u = x_2/\lambda F$
$v = y_2/\lambda F$

The transform along the y axis is Gaussian since

$$FT(\exp - a^2 y^2) = \frac{\sqrt{\pi}}{a} \exp - (\pi v/a)^2 \tag{9}$$

The transform along the x_2 or frequency axis is the key information. This transform equals the convolution of the transforms of each x_1 dependent factor. Recalling that $K = \Omega/v_a$ and that the transform of exp j a x is a delta function, the final result is:

$$\begin{aligned} FT(E_d) = {} & \gamma \exp j[(\omega - \Omega)t - \phi] \exp j\, k\, F \cos \theta_i \exp - (\frac{\pi v}{a})^2 \cdot \\ & FT\ \{S(t + \frac{x_1}{v_a})\} \otimes \exp j\left[2\pi F \frac{\tan \theta_d}{\cos \theta_d}(u - \frac{\Omega_C}{2\pi v_a})\right] \cdot \exp - \left[\frac{\pi \left(u - \frac{\Omega_C}{2\pi v_a}\right)}{a \cos \theta_d} \right]^2 \end{aligned} \tag{10}$$

where $\otimes$ is the convolution symbol.

The first part of (10) indicates that at the FT or lens back focal plan exists a frequency shifted (downshifted) light distribution. The next exponential term represents a constant phase term and is of no consequence. The v dependent term indicates (as already mentioned) that along the y_2 axis the distribution is Gaussian. The 1/e total width equals $4\lambda F/\pi D$. Within the bracket is the convolution of two distributions. The first which is the FT of $S(t+x_1/v_a)$ we ignore temporarily. The last term shows that along the x_2 axis a Gaussian light distribution resulting from the RF center frequency exists. Its centroid is located at the $x_2 = \lambda F \Omega_C / 2\pi v_a$. By inference, all pure sinsusoids in the signal result in frequency shifted Gaussian light distributions whose centroid locations vary linearly with RF frequency. Thus, the x_2 axis is the so-called frequency axis.

A multiplicative phase term which varies non-linearly with frequency also appears immediately after the convolution sign. There will, therefore, be some distortion of the RF signal. Over a 100 MHz bandwidth for a 30 mm focal length, the phase distortion is less than 2π radians and probably not significant. Over a 10 MHz bandwidth the distortion is only about 0.1 radians.

As for FT $\{S(t+x_1/v_a)\}$, if $S(t)$ is a slowly varying function over the optical beam diameter D, the overall transform expression and the overall transform profile remain Gaussian (single frequency transform). It S is rapidly varying, e.g., pulsed RF signals whose duration is much less than the acoustic travel time through D, the shape of its transform will dominate the overall transform expression. For example, if the narrow pulse is rectangular (neglecting transducer and other bandwidth limitations) then the light distribution at the FT plane will tend toward a sinc dependence.

In this analysis it has been assumed that the lens axis bisects the incident optical beam at the Bragg Cell, as in Figure 3. This assumption is implicit in (8). But for achievement of variable delay, the incident beam is translated along the x axis and is, therefore, displaced from the lens axis. This results in an extra linear phase term in (10). The term equals $\exp - (j\, 2\, s\, u)$ where s is the translated distance from lens axis, and can be shown to equal $\exp - (j\, \Omega\, t_d)$ where t_d equals the acoustic transit time between where the light intersects the acoustic beam and where the lens axis intersects the acoustic beam.

SIGNAL RECONSTRUCTION

The photomixer at the lens back focal plane is illuminated by the transformed diffracted light and a local oscillator (LO) spot. The LO light has not been frequency translated by Bragg Cell B1. The LO beam intersects the FT lens at the same location as the diffracted beams because of the action of the 2 beamsplitters and 2 mirrors. Also, after reflection from the second beamsplitter, the LO is approximately parallel to the diffracted beams. This guarantees that at the photomixer, the diffracted and LO light propagation directions will be nearly parallel. This parallelism is necessary for efficient optical heterodyning.

In particular, it can be shown that the photomixer current is proportional to $\mathrm{sinc}\,(L \sin \alpha/\lambda)$ where α equals the misalignment angle between the 2 optical beams and L equals the spot width. As an example, if L = 10 microns, then α must be less than 0.9° so that the sinc function is greater than or equal to 0.9. $\lambda = 0.63$ microns was assumed. An analysis which further describes the signal reconstruction at the photomixer is given in Ref. 2.

Insertion Loss

The insertion loss expressed in decibels is the sum of the gains (losses) at the acoustic transducer, Bragg Cell, photomixer and beamsplitters. Figure 6 illustrates the system and includes some calculated and observed insertion loss data. The beamsplitter losses are included in the Bragg Cell and photomixer interactions. The acoustic transducer gain is assumed to equal -6 dB. The gain of the Bragg Cell can be expressed as

$$G_2 = \left(\frac{\pi}{\lambda}\right)^2 \frac{M_2 L P_{LASER}}{2H} \tag{11}$$

where M_2 is its A/O figure of merit, L/H is the ratio of acoustic beam width to height and P_{LASER} equals the laser power.

The gain of the photomixer can be expressed (Ref. 3) as

$$G_3 = \left(\frac{\eta e}{h\nu}\right)^2 \frac{P_{LASER} R_{eq}}{4m} \tag{12}$$

where η = the quantum efficiency
e = the electronic charge
h = Planck's constant
ν = the optical frequency
R_{eq} = the equivalent resistance of the photo detector and output circuit
$1/m$ = the fraction of the available LO power which overlaps the diffracted light

Figure 6 shows that the insertion loss varies inversely as the square of the laser power.

System Bandwidth

The bandwidth, assuming it is not limited by the Bragg Cell, is determined by the LO distribution width along the frequency axis at the photomixer. Figure 7 illustrates this characteristic. The LO spot width is essentially diffraction limited. Therefore, by narrowing the LO beam before it is focused by the FT lens, an increasingly larger spot may be obtained. This arrangement is not shown in Figure 1.

The bandwidth equals the ratio of the local oscillator spot width to the frequency gradient along the frequency axis. For a non-beamsteered Bragg Cell transducer, assuming normal Bragg diffraction, this gradient is expressed

$$\frac{\Delta x_2}{\Delta f} = \frac{F\lambda}{v_a}\Delta f \tag{13}$$

The 1/e Gaussian spot width equals $4\lambda F/\pi D$ where D equals the local oscillator beamwidth. Therefore, the bandwidth is expressed as

$$BW = \frac{4\lambda F/\pi D}{F\lambda/v_a} = \frac{4}{\pi}\ \frac{v_a}{D} \tag{14}$$

A typical bandwidth is 8 MHz. Here we assumed v_a equals 4000 m/sec and D equals 0.6 mm.

Demonstrated Results and Further Goals

Demonstrated results include:

- 0.1 to 0.5 microsecond variable delay
- 5 MHz bandwidth (center frequency = 80 MHz)
- 70 dB insertion loss.

The immediate goals are:

- reduction of insertion loss
- increase in bandwidth
- achievement of longer delays.

A desirable long term goal is implementation using, not discrete components, but integrated ones. The benefits of so doing are well known and relate to alignment stability, size, and long term cost.

Applications

We close this paper by listing some areas where wideband variable delay is of value. These areas include:

- Adaptive Array Antenna Processing
- Radar Jamming (Range Gate Pulloff)
- Communications - Suppress Multipath Effects
- Spread Spectrum Intercept Receiver - Feature Detection
- Processing of Radar Signals.

REFERENCES

1. J. C. Sethares, J. M. Owens, and C. V. Smith, "MSW Nondispersive, Electronically Tunable Time Delay Elements," Electronics Letters, Vol. 16, No. 22, Oct. 1980, pp. 825-826.

2. J. N. Lee, N. J. Berg, M. W. Casseday, and P. S. Brody, "High-Speed Adaptive Filtering and Reconstruction of Broad-Band Signals Using Acousto-Optic Techniques," IEEE Ultrasonics Symp. 1980, pp. 488-491.

3. H. Melchior, M. Fisher, and F. Arams, "Photodetectors for Optical Communications Systems," Proc. IEEE, 58(10), Oct. 1970.

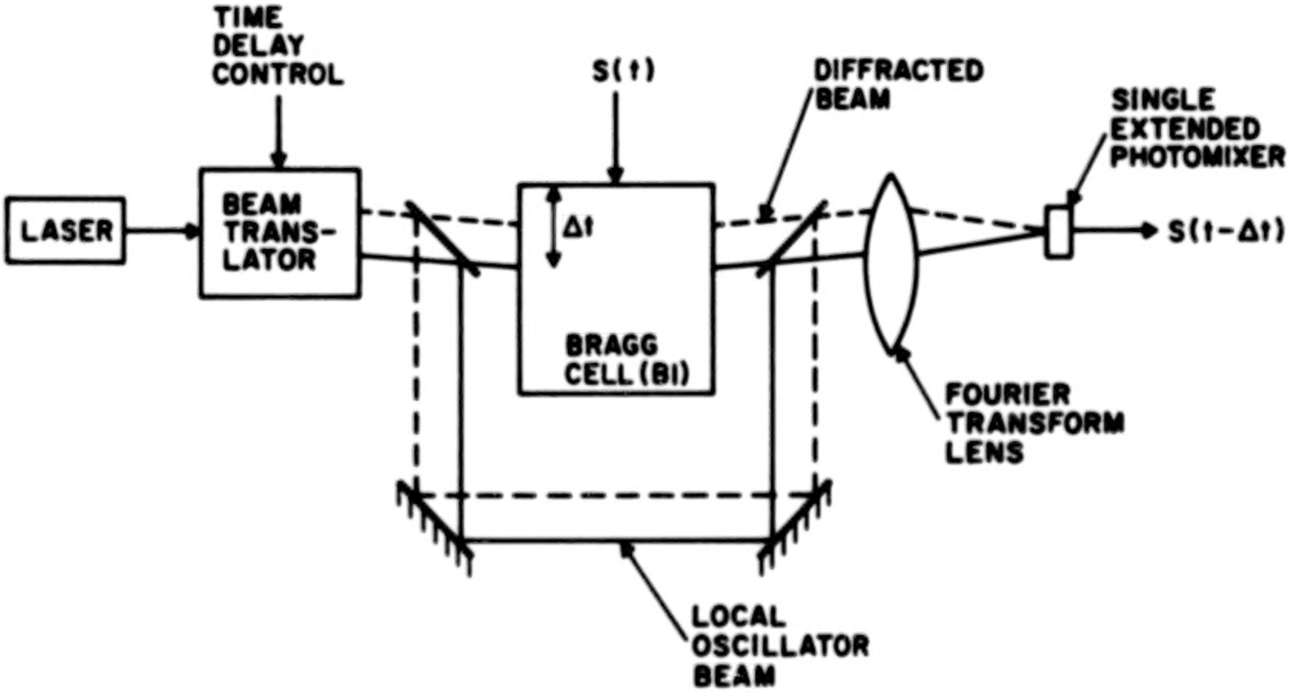

TIME DELAY CONTROL
S(t)
DIFFRACTED BEAM
SINGLE EXTENDED PHOTOMIXER
LASER
BEAM TRANS-LATOR
Δt
BRAGG CELL (BI)
S(t-Δt)
FOURIER TRANSFORM LENS
LOCAL OSCILLATOR BEAM

Figure 1.- A/O continuously variable wideband time delay.

BEST DOCUMENT AVAILABLE

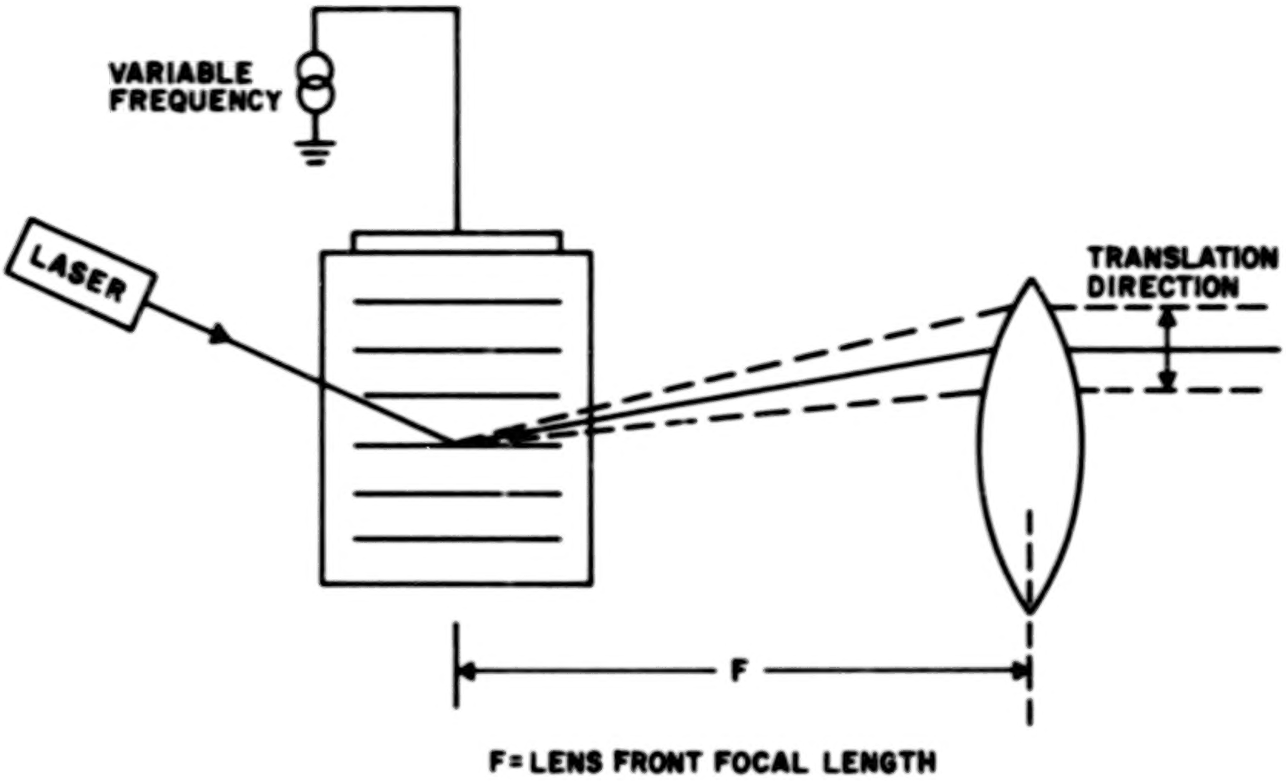

VARIABLE
FREQUENCY
LASER
TRANSLATION
DIRECTION
F
F = LENS FRONT FOCAL LENGTH

Figure 2.- Laser beam translator.

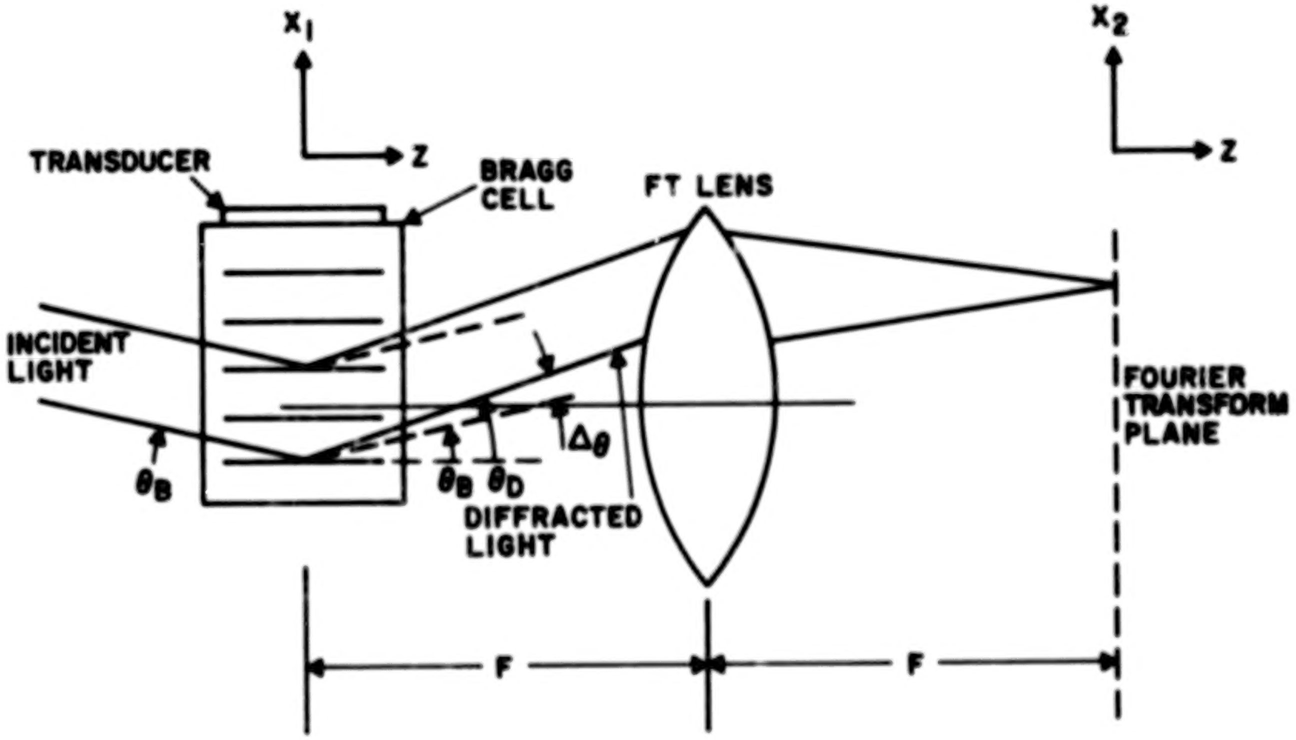

x_1
Z
x_2
Z
TRANSDUCER
BRAGG
CELL
FT LENS
INCIDENT
LIGHT
θ_B
θ_B
θ_D
Δ_θ
DIFFRACTED
LIGHT
FOURIER
TRANSFORM
PLANE
F
F

Figure 3.- Fourier transformation of Bragg diffracted light.

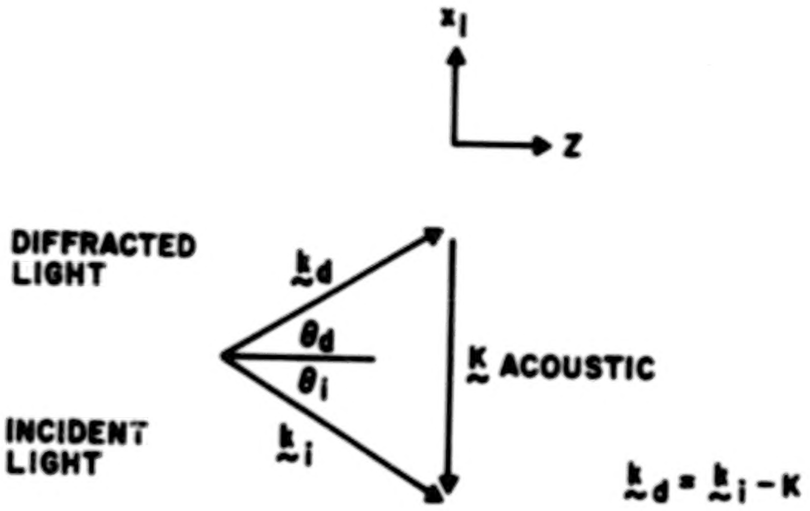

x
z
DIFFRACTED LIGHT
$\underset{\sim}{k}_d$
θ_d
θ_i
$\underset{\sim}{K}$ ACOUSTIC
INCIDENT LIGHT
$\underset{\sim}{k}_i$
$\underset{\sim}{k}_d = \underset{\sim}{k}_i - K$

Figure 4.- Bragg diffraction momentum relation.

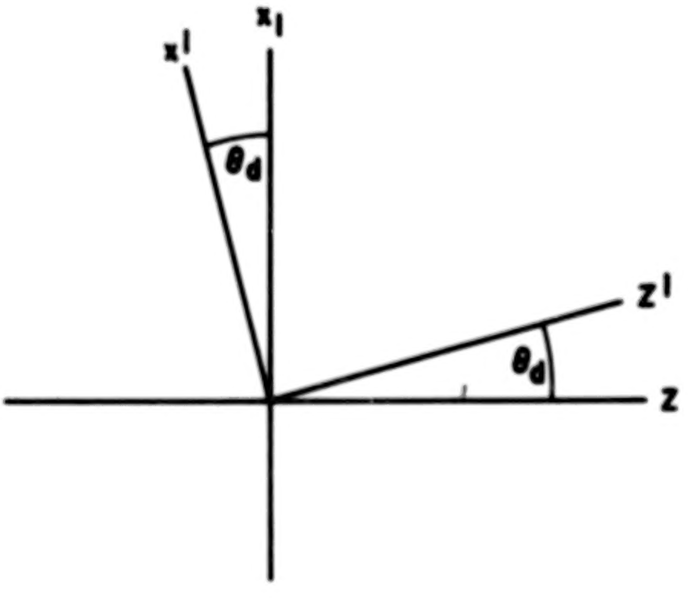
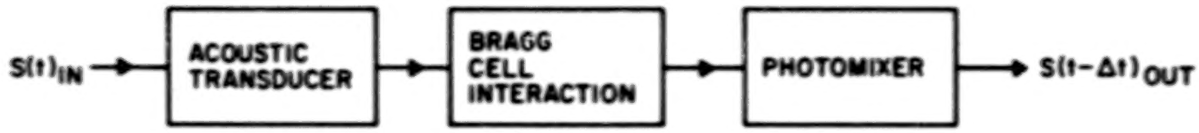

x'
x
θ_d
z'
θ_d
z

Figure 5.- Rotated XZ coordinate system.


$S(t)_{IN}$
ACOUSTIC TRANSDUCER
BRAGG CELL INTERACTION
PHOTOMIXER
$S(t-\Delta t)_{OUT}$

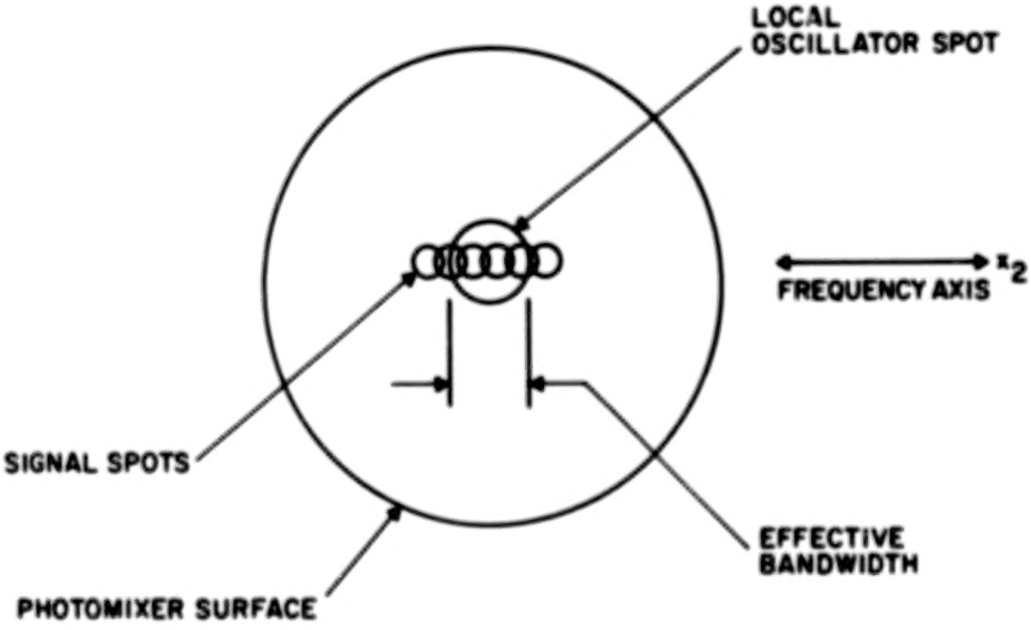
P_{LASER} (WATTS)	G_1(db)	G_2(db)	G_3^*(db)	TOTAL THEORETICAL INSERTION LOSS	OBSERVED INSERTION LOSS
.002	−6	−25	−31	−62	−70
.020	−6	−15	−21	−42	——
.20	−6	−5	−11	−22	——

$$\text{INSERTION LOSS} = G_1 + G_2 + G_3$$

$$\text{IS PROPORTIONAL TO } \left(\frac{1}{P_{LASER}}\right)^2$$

* SIGNAL BW = 10 MHz

Figure 6.- System insertion loss.


LOCAL OSCILLATOR SPOT
x_2
FREQUENCY AXIS
SIGNAL SPOTS
EFFECTIVE BANDWIDTH
PHOTOMIXER SURFACE

$$\text{BANDWIDTH} = \frac{\text{LO SPOT WIDTH}}{\Delta x_2/\Delta f}$$

Figure 7.- System bandwidth dependence.

MONOLITHICALLY INTEGRATED ACTIVE OPTICAL DEVICES*

J. Ballantyne, D. K. Wagner, B. Kushner and S. Wojtzcuk
Cornell University
Ithaca, NY 14853

SUMMARY

We discuss considerations relevant to the monolithic integration of optical detectors, lasers, and modulators with high-speed amplifiers. Some design considerations for representative subsystems in the GaAs-AlGaAs and GaInAs-InP materials systems are described. Results of a detailed numerical design of an electro-optical birefringent filter for monolithic integration with a laser diode is described, and early experimental results on monolithic integration of broadband MESFET amplifiers with photoconductive detectors are reported.

INTRODUCTION

A number of workers have been involved in the monolithic integration of detectors and lasers with amplifiers for applications in optical communication.[1] The work reported here is specifically aimed at circuits and subsystems suitable for use in very-high data rate applications. The OPFET class of photoconductive detectors in III-V compounds shows picosecond response time coupled with internal gain and low noise. In order to realize the potential of these detectors in a practical system, monolithic integration with high-speed, low-noise preamplifiers is probably a necessity. In addition, there is a need for integration of laser diodes with high-speed amplifiers in order that complete opto-electronic systems can be constructed.

Frequency tunable lasers are useful for FM communication, for frequency multiplexed systems, and offer outstanding potential for very-high data rate transmission.[2] These may be constructed by suitable combinations of gain, birefringent and electro-optical components in the laser cavity.[3,4]

Two materials systems seem especially attractive for monolithic integration of opto-electronic subsystems. They are the GaAs-AlGaAs and GaInAs-InP heterostructure systems. The first system is useful because well-characterized opto-electronic components and amplifiers have been constructed in this system, and much is known about the materials problems which must be overcome to construct successful subsystems. The GaInAs-InP heterostructure system, lattice-matched to InP substrates, is interesting because of its potential application for longer wavelength communications, where optical fibers show minimum dispersion. A further reason for interest in this system is the very high electron saturation velocity, which was shown to be the greatest for any of the GaInAsP quaternaries grown lattice-matched to InP. In previous work we showed that the electron saturation velocity in GaInAs is nearly double that of GaAs for comparable carrier densities at room temperature.[5] This enables one to construct photoconductive detectors with higher internal gain, and transistors with higher G_m and speed in the GaInAs system than in the conventional GaAs system. A drawback to the GaInAs-InP system is that unsolved materials problems must be overcome to construct successful monolithic opto-electronic circuits, and the individual devices are much less developed in this materials system, particularly the transistor technology.

---

*Portions of this work were supported by NASA Langley Research Center, The National Science Foundation, and the Joint Services Electronics Program at Cornell University. Facilities of the National Research and Resource Facility for Submicron Structures and the Materials Science Center at Cornell University were employed in this work.

SOME DESIGN CONSIDERATIONS FOR MONOLITHIC OPTO-ELECTRONIC SUBSYSTEMS

Monolithic integration of optical and amplifying devices requires semi-insulating substrates if the well-developed MESFET technology is to be utilized for amplifiers. Such substrates are readily available in both GaAs and InP; however, problems immediately arise for MESFET type transistors on InP substrates. From initial considerations, it might seem that a simple GaInAs, InP heterostructure interface would be useful for both optical and electrical confinement, and would be suitable for construction of MESFETs. One of the problems with such interfaces grown by MBE is shown in Fig. 1,[6] where it is seen that in an abrupt GaInAs-InP interface substantial leakage currents can flow which prevent construction of good transistors. In these MBE grown layers the interfacial currents were linked with the growth process for the interface, which may result in an InAsP interfacial layer with poor electrical characteristics. These problems were initially solved by incorporating Al into the system to provide an AlInAs interfacial layer between the GaInAs active region and the InP substrate.

While this approach solved problems with leakage current in transistors, it raises further problems for laser diodes. A schematic diagram of a proposed heterostructure system for monolithically integrated opto-electronic circuits on indium phosphide substrates is shown in Fig. 2. The structure indicated schematically there has the potential for producing good transistors for amplifiers, good optical detectors, and good lasers. The requirements for OPFET type detectors are similar to those of the MESFETs, which are that one be able to construct low-leakage Schottky barrier gates on active layers which are thin enough to be pinched off at a reasonable voltage, and that the structure show no excessive leakage currents through interfaces. The structure shown in Figure 2 meets these requirements. The AlInAs layer under the gate serves to make a good Schottky barrier as was previously demonstrated.[7] Fig. 2 is schematic in the sense that details of the ohmic contacts and Schottky barriers are not shown. Under the contacts, the undoped upper layers must either be removed, doped, or consumed in the contact alloying process. There are various approaches for doing these steps, the details of which are not discussed here. The requirement for pinching off the active FET layer at a reasonable gate voltage means that the thickness of the layer must be 0.2 microns or below for a doping of 10^{17}cm^{-3}, or that the doping must be reduced for thicker active layers.

The AlInAs layer between the substrate and active layer serves to reduce the leakage currents in the transistor to reasonable values. It also serves to provide excellent carrier and optical mode confinement for the laser structure. One advantage of this material system for the TJS laser structure shown is that the cladding layers can be easily made semi-insulating, providing excellent current confinement in the TJS laser diode. Another advantage of the structure shown for the TJS laser is that it incorporates two-dimensional index guiding, which should give extremely good mode confinement. The parallel metal stripes on the surface serve as a slot waveguide to confine the lasing mode in the direction parallel to the surface. We have done extensive calculations and modeling on various geometries of these slot waveguides, some of the results of which are included in the following section.

The problem with the AlInAs interlayer is that in cases studied to date, it reduces the quality (in terms of luminescence linewidth and efficiency) of the GaInAs active layer.[8] However other workers have shown that this is not a fatal limitation.[9] Substantial materials work remains to be done to solve this problem.

SLOT WAVEGUIDES FOR MONOLITHIC INTEGRATION

The slot waveguide configuration shown schematically in Fig. 2 appears to be extremely promising for monolithic integration in III-V compounds. The reasons for this are the experimental demonstration of stable guiding with very low threshold injection lasers in the AlGaAs heterostructure system[10], the compatibility with TJS lasers as shown in Fig. 2, and the obvious compatibility with electro-optic devices employing planar surface electrodes.

We have done extensive numerical calculations on the properties of slot waveguides in various geometries, and confirm that excellent low-loss waveguides can be made in various III-V compound heterostructure systems. These waveguides contain as an inherent component the electrode structures necessary for inducing electro-optic effects, or pumping TJS lasers. A further advantage of this waveguide geometry is that it is compatible with optical pumping, which is convenient for fundamental investigations. The basic geometry of the slot waveguide system we analyzed is shown in Fig. 3. The long direction of the slot is oriented along a <110> direction so that electric fields applied by the electrodes are along along <110> directions as required for usable electro-optic devices. The various guide geometries we analyzed are shown in Fig. 4. These waveguides are not only useful to provide a passive guide for a laser structure, but are also suitable for combining an electro-optically active device within a laser cavity, such as a Q-switch or a tunable birefringent filter (which will provide a frequency tunable diode laser).

We previously demonstrated the feasibility for rapid electro-optic frequency tuning of diode lasers,[3] and are now constructing a monolithically integrated version of this system. While the simple slot waveguide geometry shown as Type I can provide good guiding and low enough loss for a laser cavity, additional considerations arise when one considers electro-optic devices. (See Fig. 5 for the results in the simple system). Two additional considerations of importance for electro-optic devices are achievement of phase matching between TE and TM modes so that good conversion efficiency can be obtained on application of an electric field, and the requirement for some electro-optic structures that both TE and TM modes exhibit low loss. These requirements led to the investigation of the other waveguide geometries shown as Types II through VI. The only geometries which meet both of these requirements, of good phase matching and low loss for TE and TM modes, are Types V and VI, with Type VI being the preferred guide. As shown in Figs. 6 and 7, both of these systems can provide good phase matching; however, the Type VI guide appears to have relaxed fabrication tolerances as compared to Type V. The fabrication tolerances are very sensitive functions of the Aluminum content in the cladding layers. Decreasing Al content relaxes these tolerances. The mode confinement for a representative Type VI guide is shown in Fig. 8 and the electro-optic coupling efficiency for TE to TM modes is shown in Fig. 9. As shown in Fig. 5, the simpler waveguide types I-IV show a great difference between loss for TE and TM modes. Such a structure would be particularly useful for Q-switching a diode laser in a monolithically integrated configuration.

INTEGRATED SUBSYSTEMS

In Figs. 10 and 11 we show the circuit diagram and photograph of a finished circuit of a monolithically integrated OPFET detector and three stage wide bandwidth amplifier. This system has been designed to give a voltage gain of 7, a power gain of 250 (over 23 dB), and a frequency response exceeding 5 GHz. One unique feature of this chip is the inclusion of a new type of interdigitated OPFET detector shown in Fig. 12. The areal response of this detector is shown in Fig. 13 which is the photocurrent due to a HeNe laser spot, raster scanned over the detector. The low frequency gain measured in this experiment was 14, which is most likely a factor of 2 larger than would be measured at high speeds. Such high speed measurements are now in progress. In Figs. 14 and 15 we show the circuit diagram and layout for an integrated laser/driver-transistor circuit being fabricated in the materials structure shown in Fig. 2. The two small FETs connected in series on the gate of the driver transistor provide a convenient, area-and-power-efficient method for setting the bias level on the laser diode.

CONCLUSIONS

In this paper we have identified what we consider to be viable approaches for the monolithic integration of detectors, lasers, amplifiers, and electro-optic components in two material systems which operate in quite different wavelength regions. For the first time, we have carried out extensive numerical calculations on slot waveguide structures in III-V compounds, and have computed parameters which are useful for the design of guides for electro-optic components in the systems. The results of these computations give information on the confinement of various geometry guides, the attenuation and the electro-optic coupling parameters of these guides. Finally, we have fabricated the most complex monolithic integrated opto-electronic circuits so far reported in III-V compounds, and have verified their performance at low frequency. We expect these circuits to give good operation at gigahertz speeds.

REFERENCES

1. D. Wilt, N. Bar-Chaim, S. Margalit, I. Ury, M. Yust and A. Yariv, IEEE J. Quant. Electr. QE-16, 390 (1980); C. P. Lee, S. Margalit, and A. Yariv, Appl. Phys. Lett. 31, 281 (1977); T. Fukuzawa, M. Nakamura, M. Hirao, T. Kuroda, and J. Umeda, Appl. Phys. Lett. 36, 181 (1980).

2. A. Olsson and C. L. Tang, IEEE J. Quant. Elect. QE-15, 1085 (1979).

3. C. L. Tang, V. G. Kreismanis and J. M. Ballantyne, Appl. Phys. Lett. 30, 113 (1977).

4. F. Reinhardt and R. A. Logan, Appl. Phys. Lett. 36, 956 (1980).

5. J. C. Gammel, H. Ohno, and J. M. Ballantyne, IEEE J. Quantum Elect. QE-17, 269 (1981).

6. J. C. Gammel, "Short Transit Time Photoconductive Detectors for High Speed Optical Communication", PhD Thesis, Cornell University, August 1980. Materials Science Center Report # 4282.

7. H. Ohno, J. Barnard, C.E.C. Wood and L. F. Eastman, Elec. Device Lett. EDL-1, 154 (1980).

8. G. Wicks, W. I. Wang, C.E.C. Wood and L. F. Eastman, "Photoluminescence Studies of III-V Ternary Alloys for Microwave Devices Grown by MBE", WOCSEMMAD '81, New Orleans, LA (Feb. 1981).

9. W. T. Tsang, J. Appl. Phys. 52, 3861 (1981).

10. Amann and Stegmueller, Integrated and Guided-Wave Optics Tech. Digest, IEEE Cat. No. 80 CH1489-4QEA, Jan. (1980).

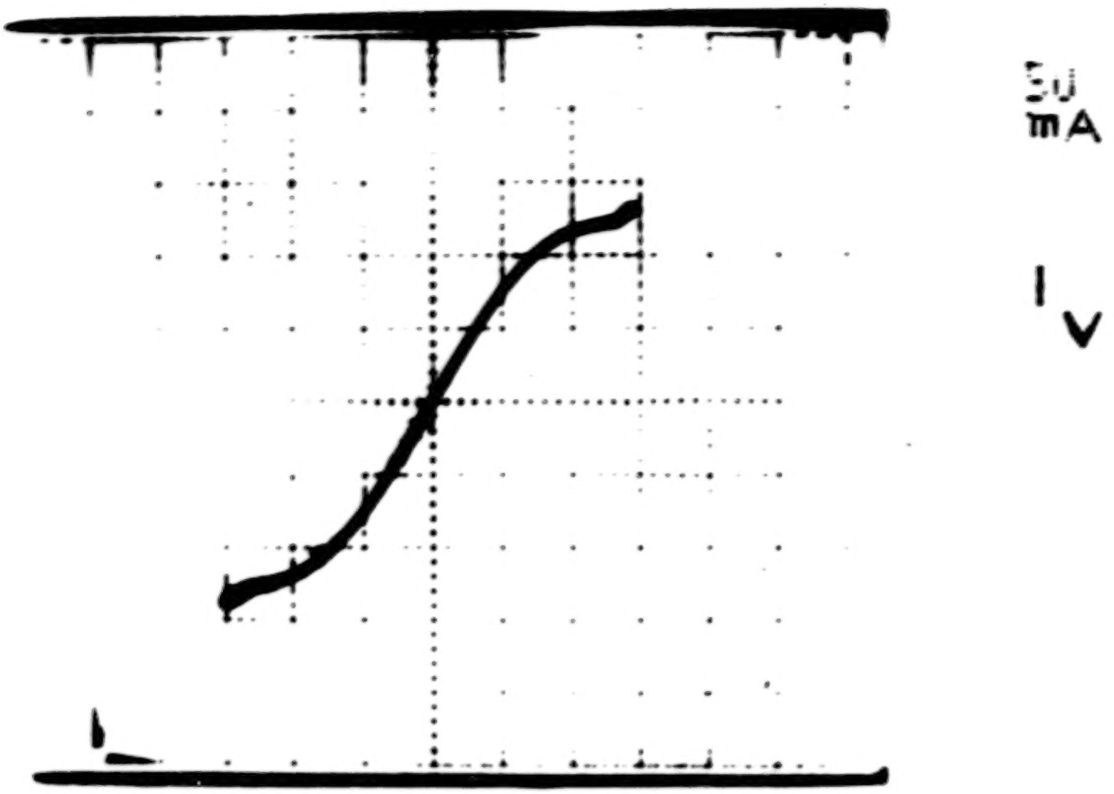


Fig. 1 I(V) curve of a GaInAs OPFET grown on an InP substrate by MBE[6]. Rise in current at end of trace is due to leakage through the substrate-epilayer interface.

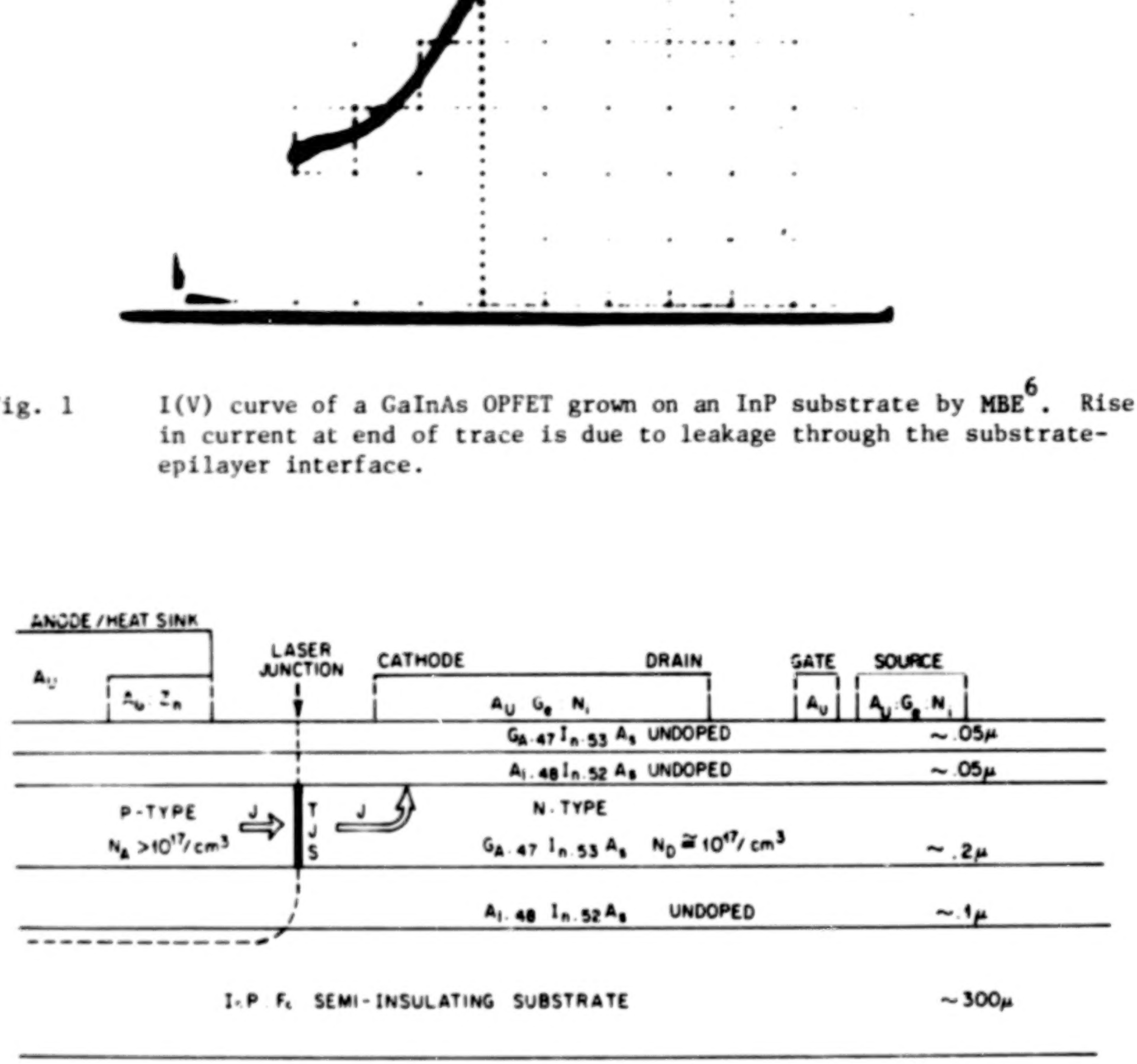
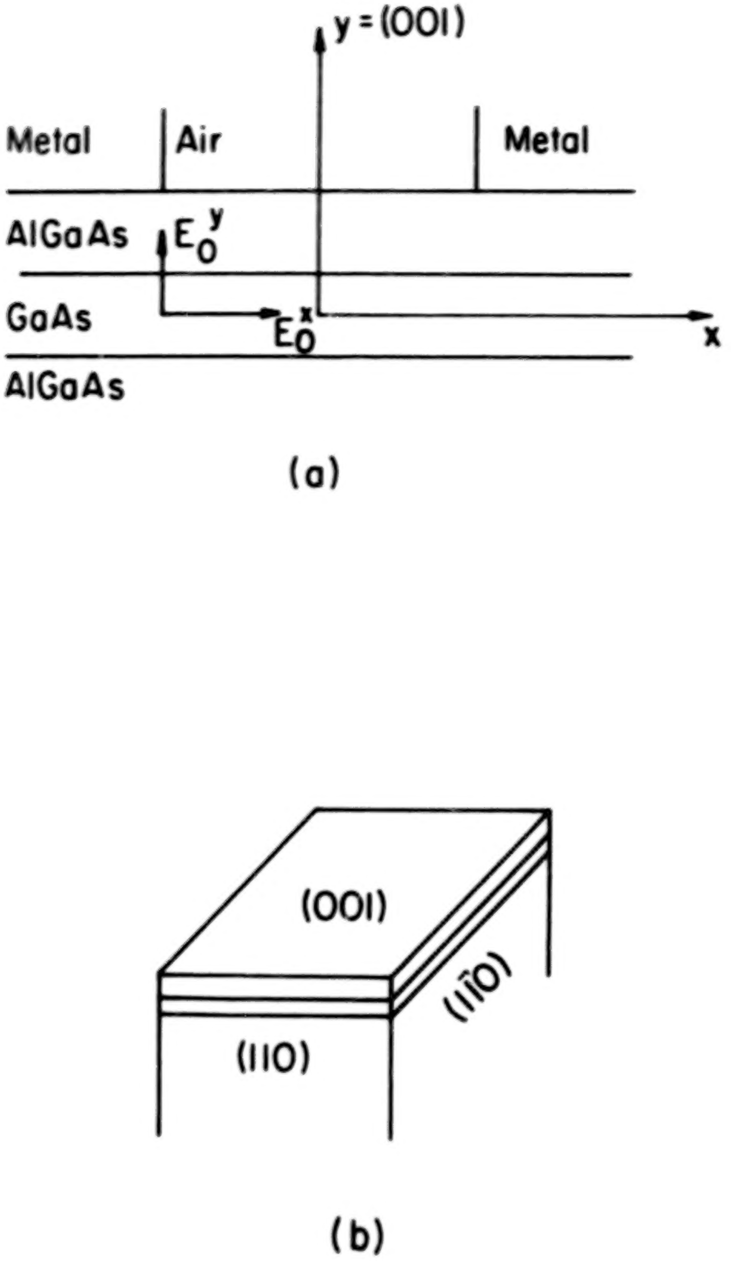

ANODE / HEAT SINK
AU
AU·ZN
LASER JUNCTION
CATHODE
DRAIN
GATE
SOURCE
AU GE NI
AU
AU·GE·NI
GA.47 IN.53 AS UNDOPED
~.05μ
AL.48 IN.52 AS UNDOPED
~.05μ
P-TYPE
NA >10¹⁷/cm³
J
T
J
S
J
N-TYPE
GA.47 IN.53 AS ND ≅ 10¹⁷/cm³
~.2μ
AL.48 IN.52 AS UNDOPED
~.1μ
I.P F. SEMI-INSULATING SUBSTRATE
~300μ

Fig. 2 Schematic diagram of proposed heterostructure system for monolithically integrated opto-electronic circuits on InP substrates.


y = (001)
Metal
Air
Metal
AlGaAs
E_0^y
GaAs
E_0^x
x
AlGaAs
(001)
(1$\bar{1}$0)
(110)

(a)

(b)

Fig. 3 Basic geometry of a slot-waveguide structure compatible with monolithic electro-optic devices in III-V compounds.

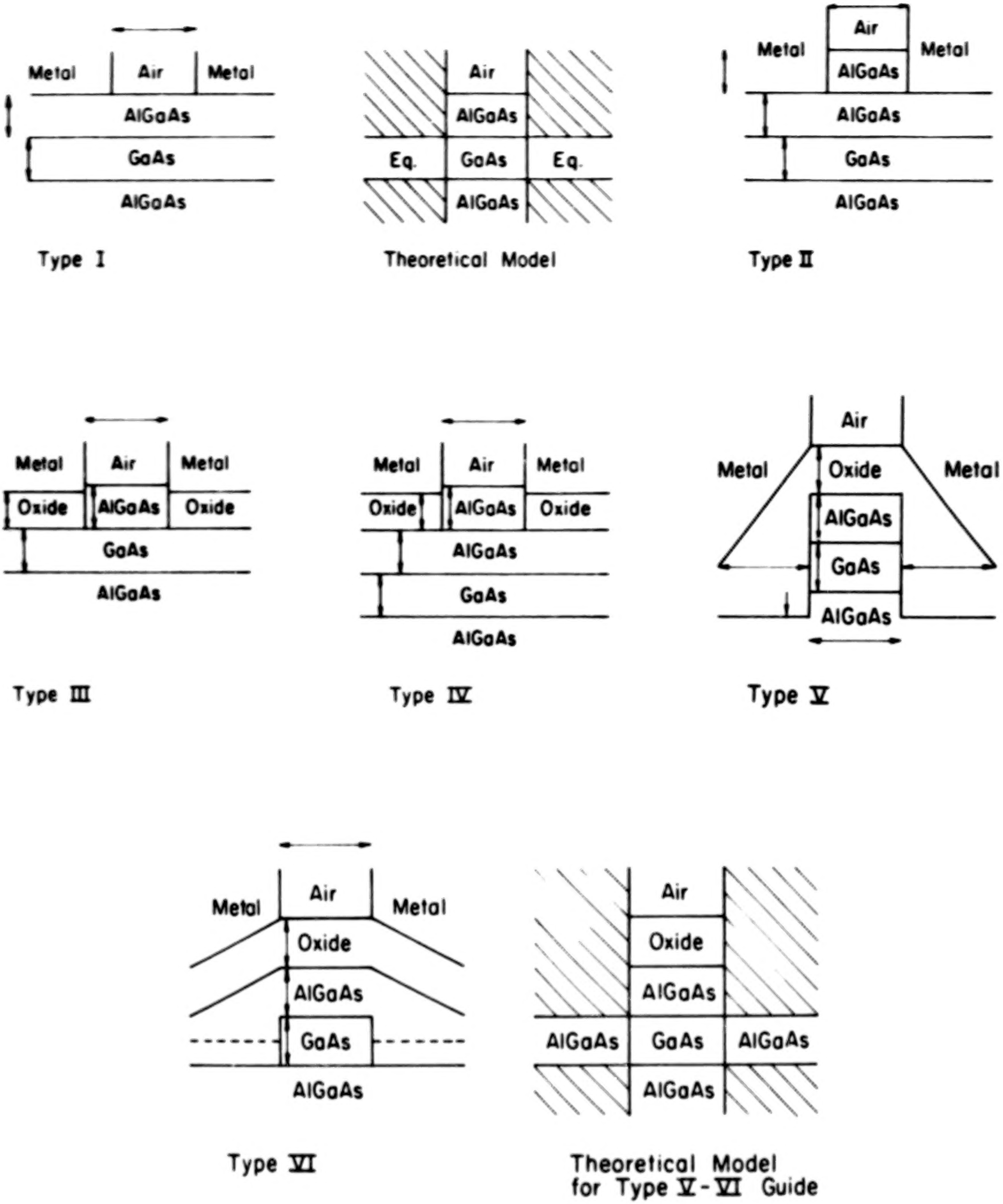
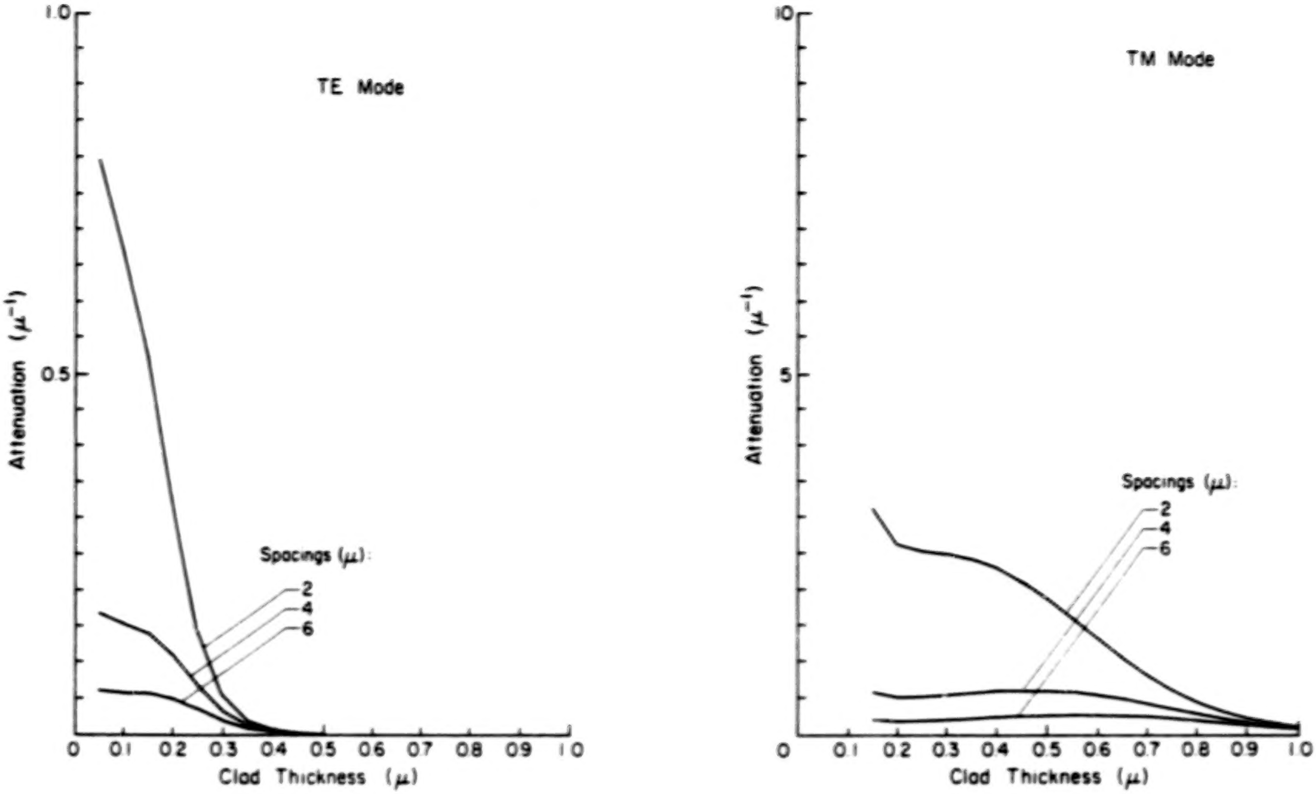
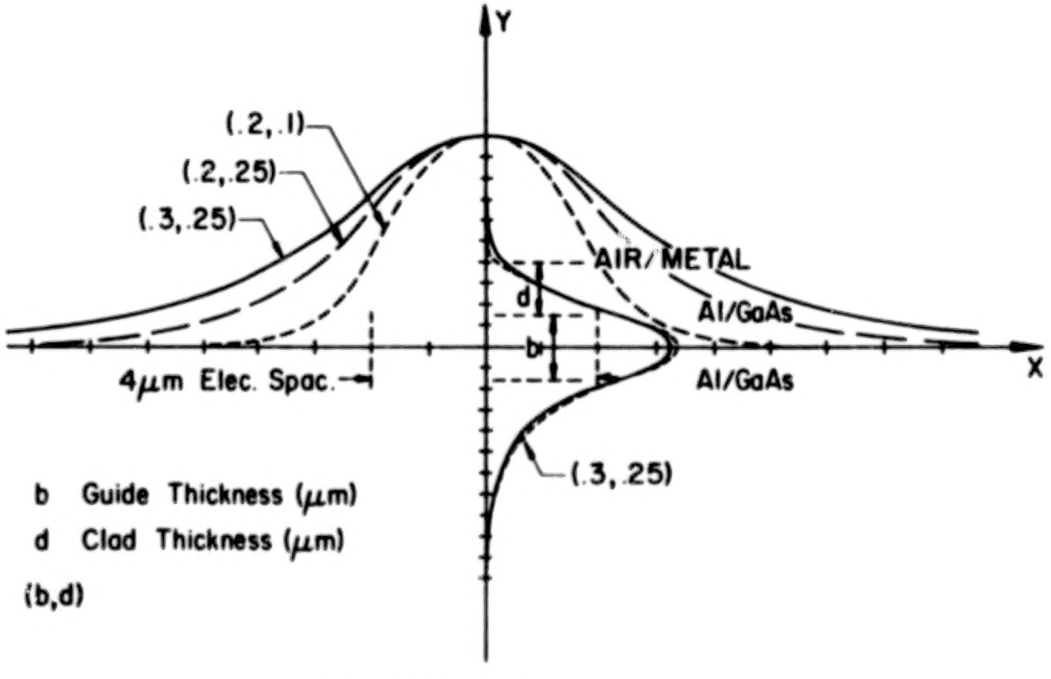

Metal
Air
Metal
AlGaAs
GaAs
AlGaAs
Type I
Air
AlGaAs
Eq.
GaAs
Eq.
AlGaAs
Theoretical Model
Air
Metal
AlGaAs
Metal
AlGaAs
GaAs
AlGaAs
Type II
Metal
Air
Metal
Oxide
AlGaAs
Oxide
GaAs
AlGaAs
Type III
Metal
Air
Metal
Oxide
AlGaAs
Oxide
AlGaAs
GaAs
AlGaAs
Type IV
Air
Metal
Oxide
Metal
AlGaAs
GaAs
AlGaAs
Type V
Metal
Air
Metal
Oxide
AlGaAs
GaAs
AlGaAs
Type VI
Air
Oxide
AlGaAs
AlGaAs
GaAs
AlGaAs
AlGaAs
Theoretical Model
for Type V - VI Guide

Fig. 4 Slot-waveguide geometries for which analysis was carried out, along with basic models used for two classes of guides.



a) Attenuation vs. clad thickness for several electrode spacings.


Y
X
(.2,.1)
(.2,.25)
(.3,.25)
AIR/METAL
d
b
Al/GaAs
Al/GaAs
4μm Elec. Spac.
(.3,.25)
b Guide Thickness (μm)
d Clad Thickness (μm)
(b,d)

b) Mode confinement.

Fig. 5 Results for modes in Type I waveguide.

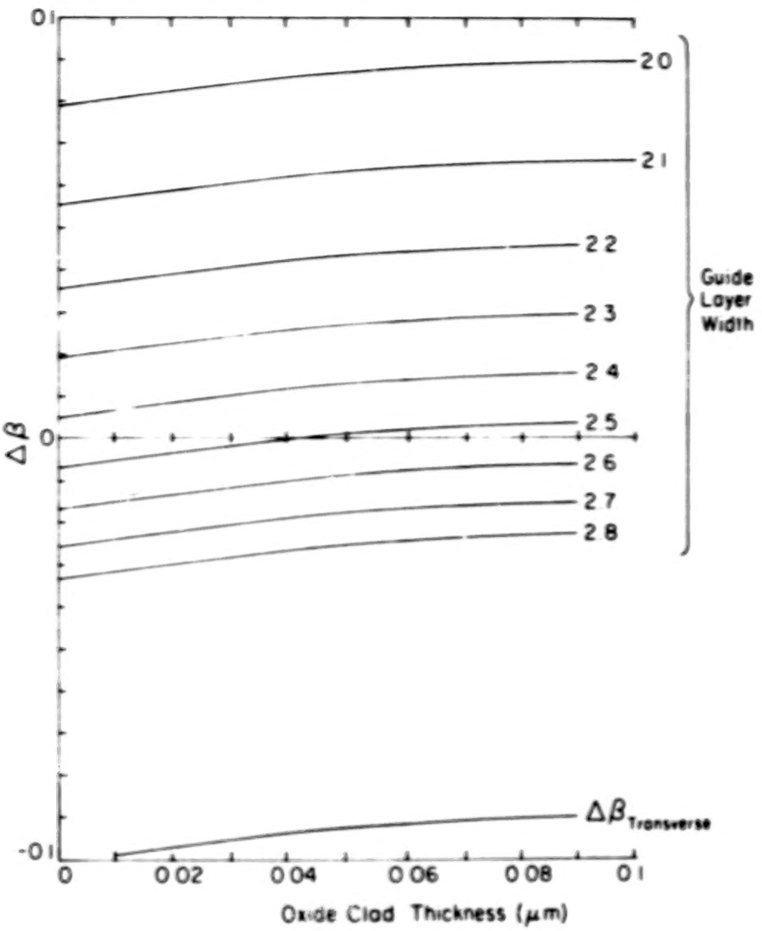


Fig. 6 Phase difference, Δβ, between TE and TM modes in Type V waveguide. Δβ is zero for a 2.5 μm wide guide and an oxide clad of 0.04 μm. Note sensitivity to fabrication tolerances.

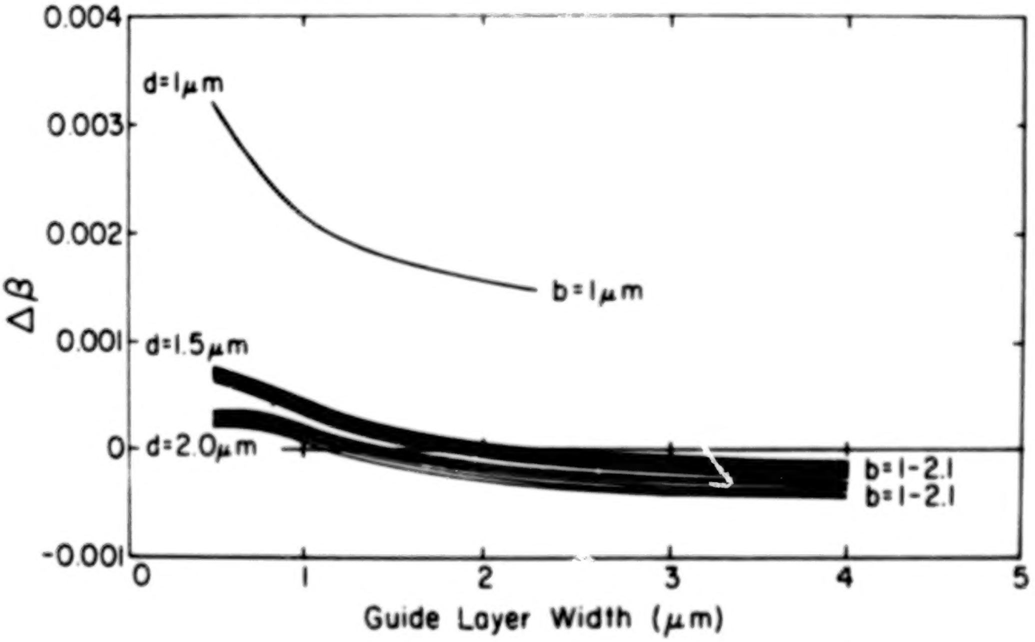


Fig. 7. Phase difference, Δβ, between TE and TM modes in Type VI waveguide: Oxide and $Al_xGa_{1-x}As$ thickness are d and b respectively. Case illustrated is for x = .05. Note large tolerances in d, b, and width to allow Δβ to be nearly zero.

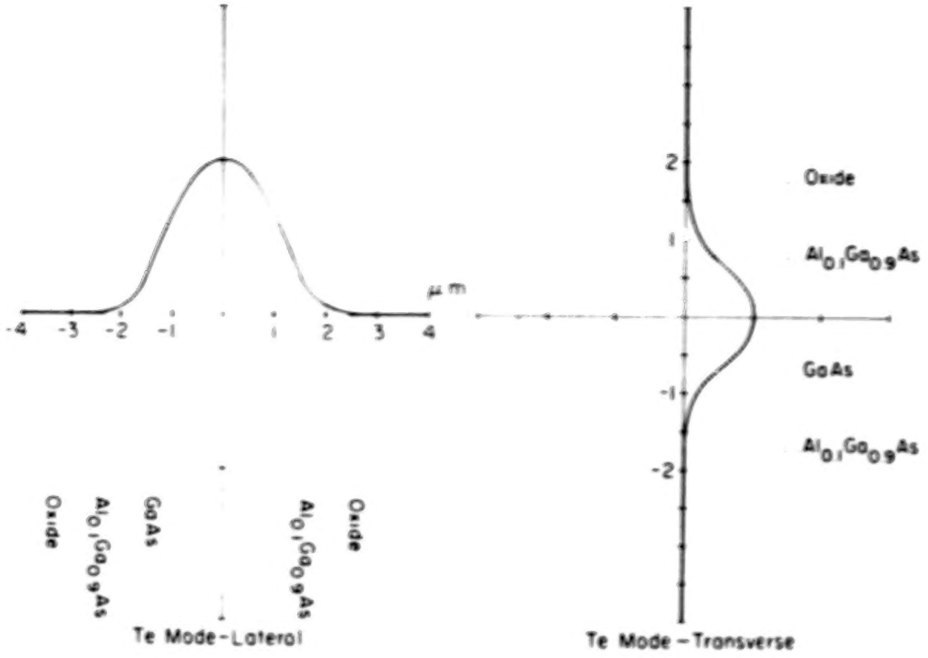


Fig. 8. Mode confinement for a representative Type VI **waveguide. Amplitudes** are normalized.

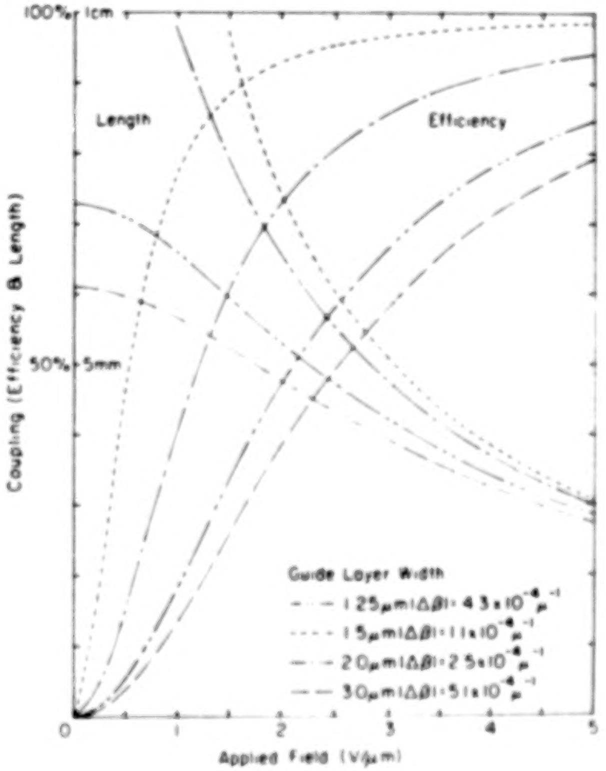


Fig. 9. Efficiency for coupling TE to TM modes in Type VI guides via the **electro-optic effect.**

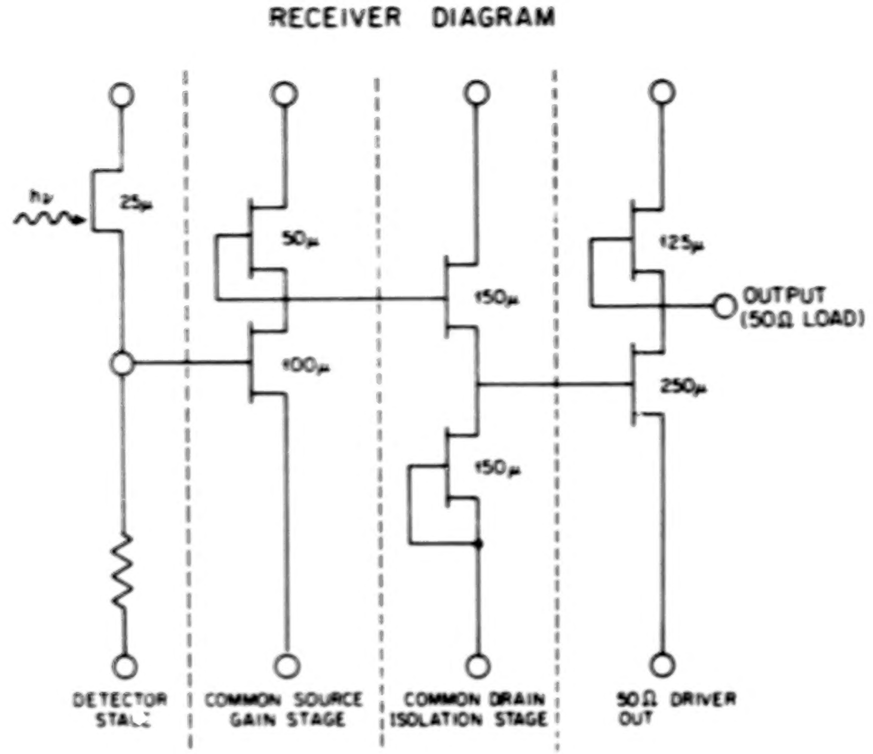

RECEIVER DIAGRAM
hν
25μ
50μ
150μ
100μ
150μ
125μ
OUTPUT (50Ω LOAD)
250μ
DETECTOR STAGE
COMMON SOURCE GAIN STAGE
COMMON DRAIN ISOLATION STAGE
50Ω DRIVER OUT

Fig. 10 Circuit diagram of a monolithically integrated receiver combining an OPFET with a 3-stage amplifier.

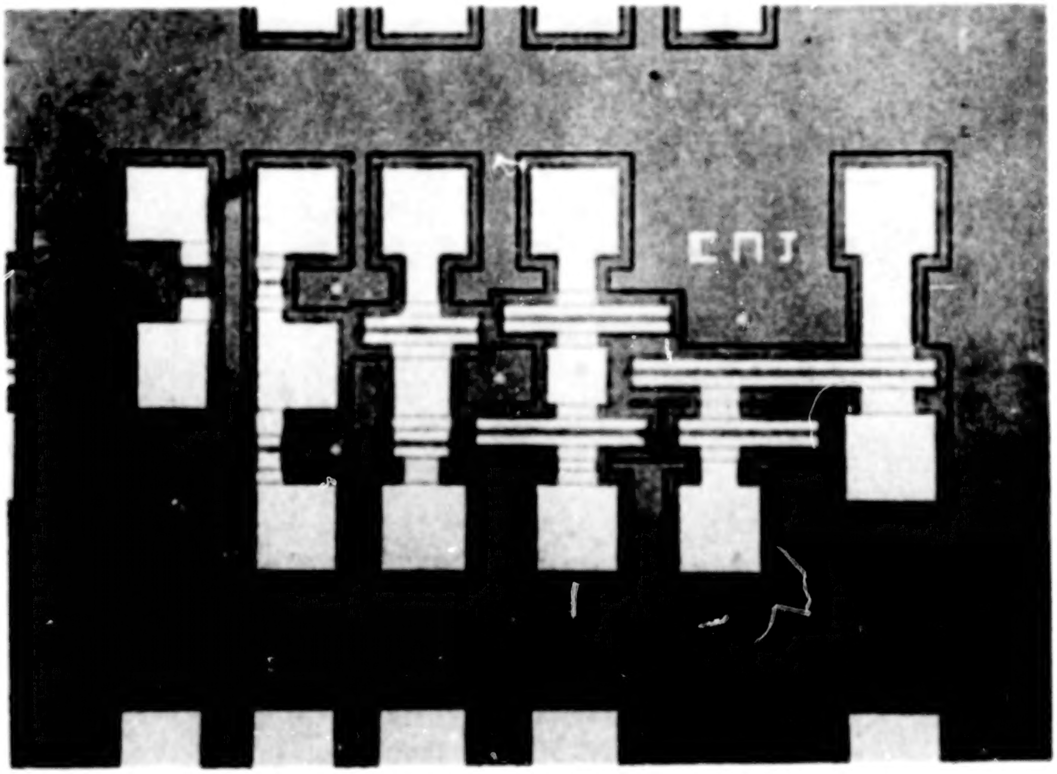
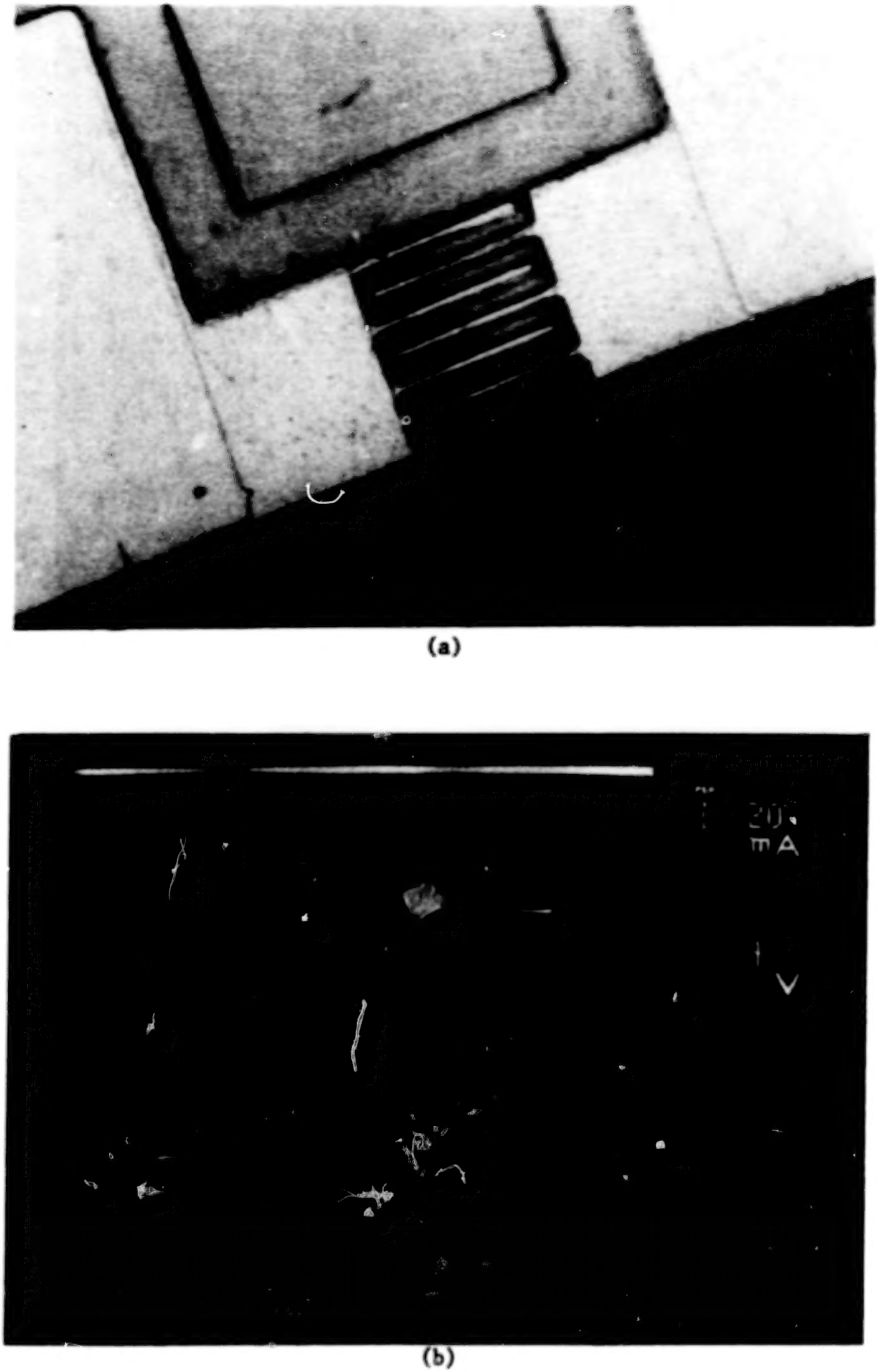


Fig. 11 Photograph of receiver of Fig. 10. Gates are missing in this photograph.



(a)

(b)

Fig. 12 Interdigitated OPFET with grooves between fingers (a), and I(V) curve of detector (b).

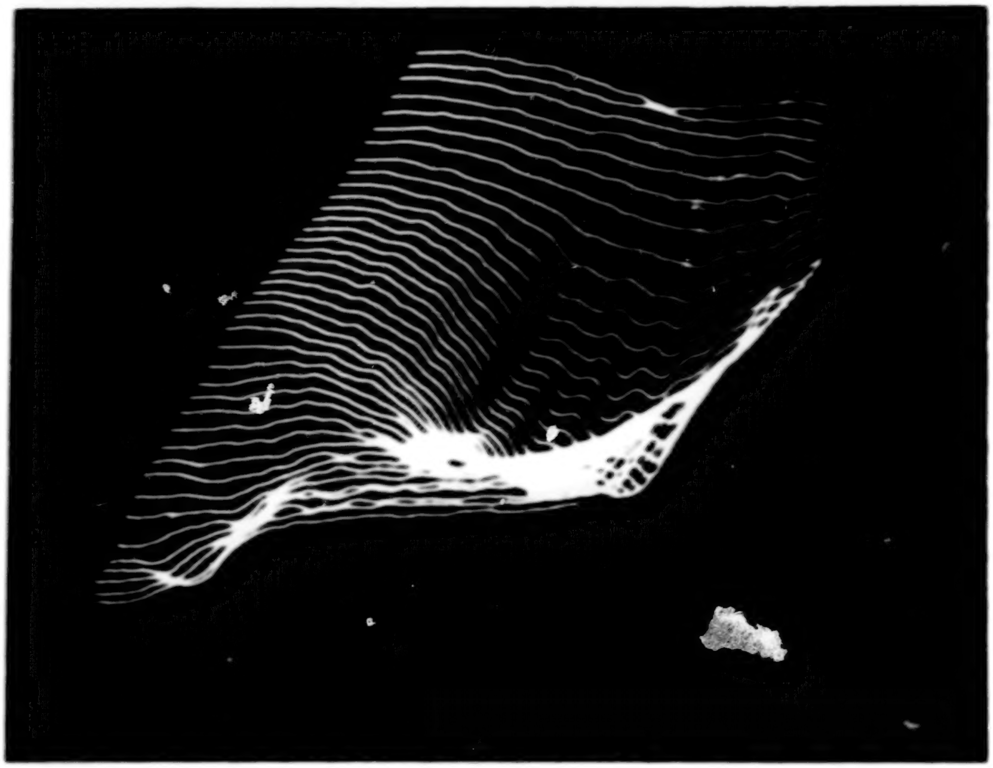


Fig. 13 Spatial response of detector in Fig. 12. Increasing response is downward. Response in lower left is an artifact due to light reflecting from bonding wire.

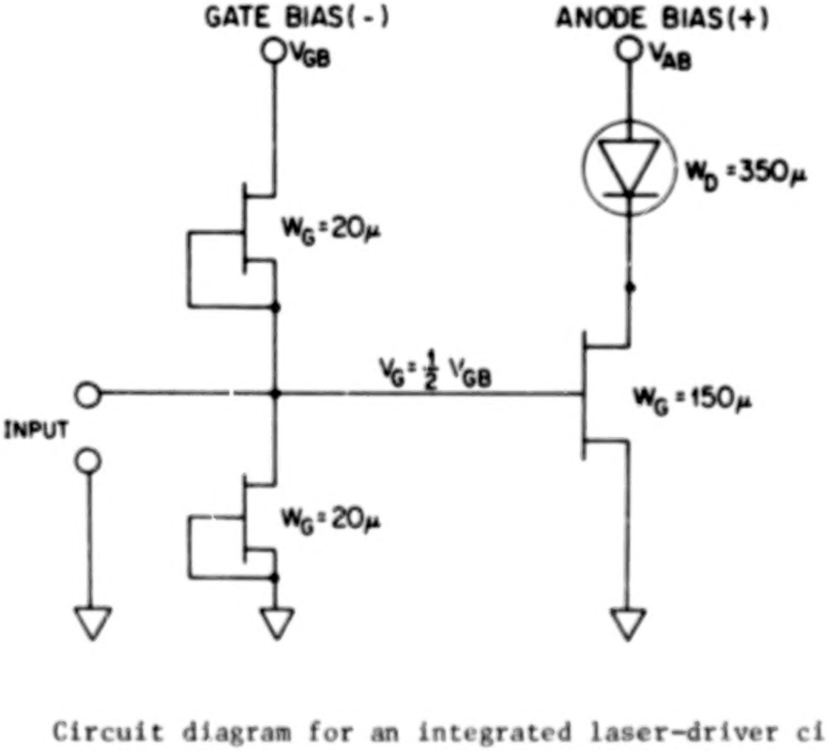

GATE BIAS(-)
V_{GB}
ANODE BIAS(+)
V_{AB}
$W_D = 350\mu$
$W_G = 20\mu$
$V_G = \frac{1}{2} V_{GB}$
$W_G = 150\mu$
INPUT
$W_G = 20\mu$

Fig. 14 Circuit diagram for an integrated laser-driver circuit being constructed in materials systems of Fig. 2.

BEST DOCUMENT AVAILABLE

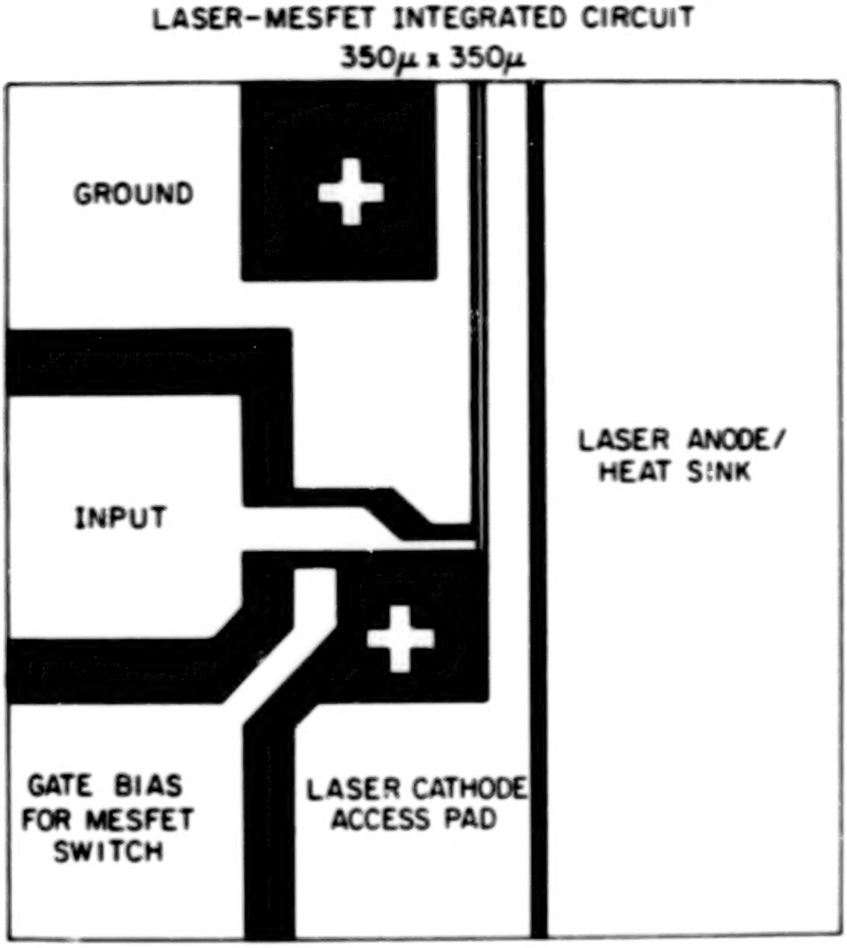

LASER-MESFET INTEGRATED CIRCUIT
350μ x 350μ
GROUND
INPUT
LASER ANODE/
HEAT SINK
GATE BIAS
FOR MESFET
SWITCH
LASER CATHODE
ACCESS PAD

Fig. 15 Layout of circuit of Fig. 14.

MICROCHANNEL SPATIAL LIGHT MODULATOR

Cardinal Warde

Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA

SUMMARY

The Microchannel Spatial Light Modulator (MSLM) is a versatile, highly-sensitive, and optically addressed device that is being developed for real-time optical information processing. The MSLM operates by converting an input optical image into a charge distribution at the surface of an electro-optic crystal. The charge distribution generates an electric field which modulates the refractive index of the crystal and thereby the phase or intensity of an image readout beam. Prototype devices employing 250 μm thick crystals have exhibited a spatial resolution of 5 cycles/mm at 50% contrast, an exposure sensitivity of 2.2nJ/cm^2 and framing rates of 40Hz with full modulation depth. The image processing operations that have been achieved using the internal processing mode of the MSLM include contrast reversal, contrast enhancement, edge enhancement, image addition and subtraction, analog and digital intensity thresholding, and binary-level logic operations such as AND, OR, EXCLUSIVE OR, and NOR.

INTRODUCTION

The Microchannel Spatial Light Modulator (MSLM)[1] is a relatively new, versatile, highly sensitive, and optically addressed device that is being developed for low-light-level real-time optical information processing. In addition to real-time operation, it can be designed to exhibit long term (months) optical information storage simultaneously.

A variety of sophisticated image processing operations can be achieved using the internal processing mode of the MSLM. These include contrast reversal, contrast enhancement, edge enhancement, image addition and subtraction, analog and digital intensity level thresholding, as well as binary-level logic operations such as AND, OR, EXCLUSIVE OR, and NOR. This paper presents an update of its characteristics.

The MSLM is illustrated in Fig. 1. It consists of a photocathode, a microchannel plate (MCP), a planar acceleration grid (optional), and an electro-optic plate in a proximity focussed configuration. The electro-optic plate generally carries a high-resistivity dielectric mirror on one side and a transparent conducting electrode on the other. The MCP is an array of small (~10 μm) semiconducting-glass pores, each functioning as a continuous dynode electron multiplier[2,3]. Spatially uniform gains (±5% variation) of 10^4 with a single MCP or 10^7 with two cascaded MCPs (chevron configuration) are typical.[4]

In operation, the write light (coherent or incoherent) incident on the photocathode creates an electron image which is amplified by the MCP and proximity focused onto the surface of the electro-optic plate. The resulting surface charge distribution creates a spatially varying electric field which modulates the refractive index of the electro-optic plate (crystal). Thus the readout light, which makes a double pass through the crystal, is spatially phase modulated.

It is the wide choice of photocathodes and electro-optic materials that gives rise to the spectral versatility of the MSLM and permits its characteristics to be tailored for specific applications. For example, the wavelength of the write signal is limited only by the spectral response of existing photocathodes. No photocathode is needed in the extreme ultraviolet and X-ray regions since the MCP responds directly to this radiation[5,6]. The readout light is limited by the spectral transmittance of electro-optic materials (the range between the middle ultraviolet and the middle infrared).

Because of its large electron gain and ability to integrate charge, the MSLM is capable of achieving the high photon sensitivities of image intensifier tubes. Devices employing one MCP are expected to be able to record visible images with intensities as low as 10^{-14} W/cm^2, the primary limitation being MCP dark current.

PRINCIPLES OF OPERATION

The dynamics of the write and erase processes in the MSLM can be understood from a study of the electron current in the gap between the grid and the crystal. The gap current density J_g depends on the voltage V_g at the crystal surface, the secondary emission characteristics of the crystal surface and on the MCP output electron energy distribution.

Figure 2 summarizes the dependence of the gap current on the above-mentioned parameters. The parameter $\bar{\delta}'(V_k)$ is an effective secondary electron emission coefficient (for electrons removed from the crystal) that takes the MCP energy distribution and the collection geometry into account. The point E in Figure 2 is an equilibrium point, and M is the point at which electrons can be removed from the crystal at the fastest rate. If the device is driven into lockout regime by lowering V_b too rapidly, the crystal surface repels the primary electrons from the MCP and it becomes very difficult to alter the state of the device.

The MSLM is therefore operated so that it is either at the equilibrium point E in Fig. 2, or it is moving along one of the trajectories of Fig. 2 toward equilibrium. For materials exhibiting the linear electro-optic effect, two write modes are possible. In the electron deposition write mode, V_b initially divides capacitively between the crystal and the gap, thereby raising V_g into the region for which $\bar{\delta}' < 1$. Then when the write light is turned on, electrons are deposited on the crystal, V_g begins to fall toward its equilibrium value V_{eq}, and the crystal voltage V_x rises. Electron accumulation may be terminated when the desired V_x is reached by turning off the write light. The crystal may be erased via secondary electron emission by flooding the photocathode with light and lowering V_g (by lowering V_b) into the region where $\bar{\delta}' > 1$. In practice, V_b is ramped downward no faster than V_g can be restored to V_{eq} by electron removal, so as to avoid lockout.

In the secondary electron emission write mode, with the write light on, V_b is ramped downward at a rate $\dot{V}_b$ chosen so that lockout is avoided and the appropriate electron distribution is removed from the crystal. The secondary emission write process terminates when either the write light is turned off, or V_b reaches its final value and V_g returns to its equilibrium value. To erase the information, V_b can be increased abruptly and the photocathode flooded with light so that the crystal gains electrons.

Optical information that has been written into the MSLM may be read out either interferometrically with schlieren optics or with crossed polarizers. The possible readout modes depend on the symmetry and cut of the electro-optic crystal. The readout mode that is used in the prototype devices employs the interference between the front and back surface reflections from a plane-parallel electro-optic crystal. If dielectric mirrors are deposited on both sides of such a plane-parallel crystal to form a Fabry-Perot interferometer of finesse F, a complete off-to-on transition can be achieved with approximately 1/F times the halfwave surface charge density of the crystal.

The MSLM can be either framed through a discrete sequence of write, read, and erase periods or operated in a continuous mode. In the framed mode, the optical write signal is integrated and stored in the form of a charge distribution; after the modulation is read out, the charge distribution is removed. In the continuous mode of operation, which can be achieved with low resistivity crystals, the instantaneous modulation is proportional to the write image intensity for all temporal variations within the bandwidth of the system.

CHARACTERISTICS OF PROTOTYPE DEVICES

The most recent vacuum-demountable prototype devices built at M.I.T. employed 0.25 mm - thick oblique-cut $LiNbO_3$ crystals coated on one side with a dielectric mirror and on the other with a transparent indium-tin-oxide electrode. A 25-mm diameter MCP was used in these devices, and the metal electrode on the input face of the MCP served as the photocathode for the 254 nm write light.

The spatial resolution of these devices was approximately 5 cycles/mm at 50% contrast, and image degradation after several hours of storage time was negligible. Figure 3 shows how the spatial resolution of the MSLM depends on device geometry, and Fig. 4 shows a sample of some of the image processing operations that have been achieved with a vacuum demountable device. The sequence of steps for achieving these operations are described in Ref. 1.

In an earlier vacuum demountable device employing a Z-cut $LiNbO_3$ Fabry Perot crystal, a framing rate of 40Hz at full modulation depth was achieved. Additionally, we have observed an exposure sensitivity of 2.2 nJ/cm^2 at 655 nm wavelength in a vacuum sealed device employing a bialkali photocathode and a Z-cut $LiTaO_3$ crystal 0.5 mm thick. This device exhibited extremely long image storage times as shown by the history of image degradation shown in Fig. 5.

COMMENTS

The resolution characteristics of the MSLM have been improved by reducing the thickness of the crystal. Further improvements in resolution are expected in the near future. The large class of image processing operations that can be achieved by internal processing and the demonstrated potential for kilohertz framing rates are expected to make the MSLM useful for a large number of signal processing operations.

REFERENCES

1. C. Warde, A. M. Weiss, A. D. Fisher and J. I. Thackara, "Optical Information Processing Characteristics of the Microchannel Spatial Light Modulator," Applied Optics 20, 2066-2074, 1981.
2. B. Leskovar, Phys. Today 30, 42, No. 11, Nov. (1977).
3. J. L. Wisa, Nucl. Instrum. Methods 162, 587 (1979).
4. Applications for Microchannel Plates, Varian Associates, Palo Alto, Calif.
5. P. J. Bjorkholm, L. P. VanSpeybroeck, and M. Hecht, Proc. Soc. Photo-Opt. Instrum. Eng. 108, 189 (1977).
6. K. W. Dolan and J. Chang, Proc. Soc. Photo-Opt. Instrum. Eng. 106, 178 (1977).

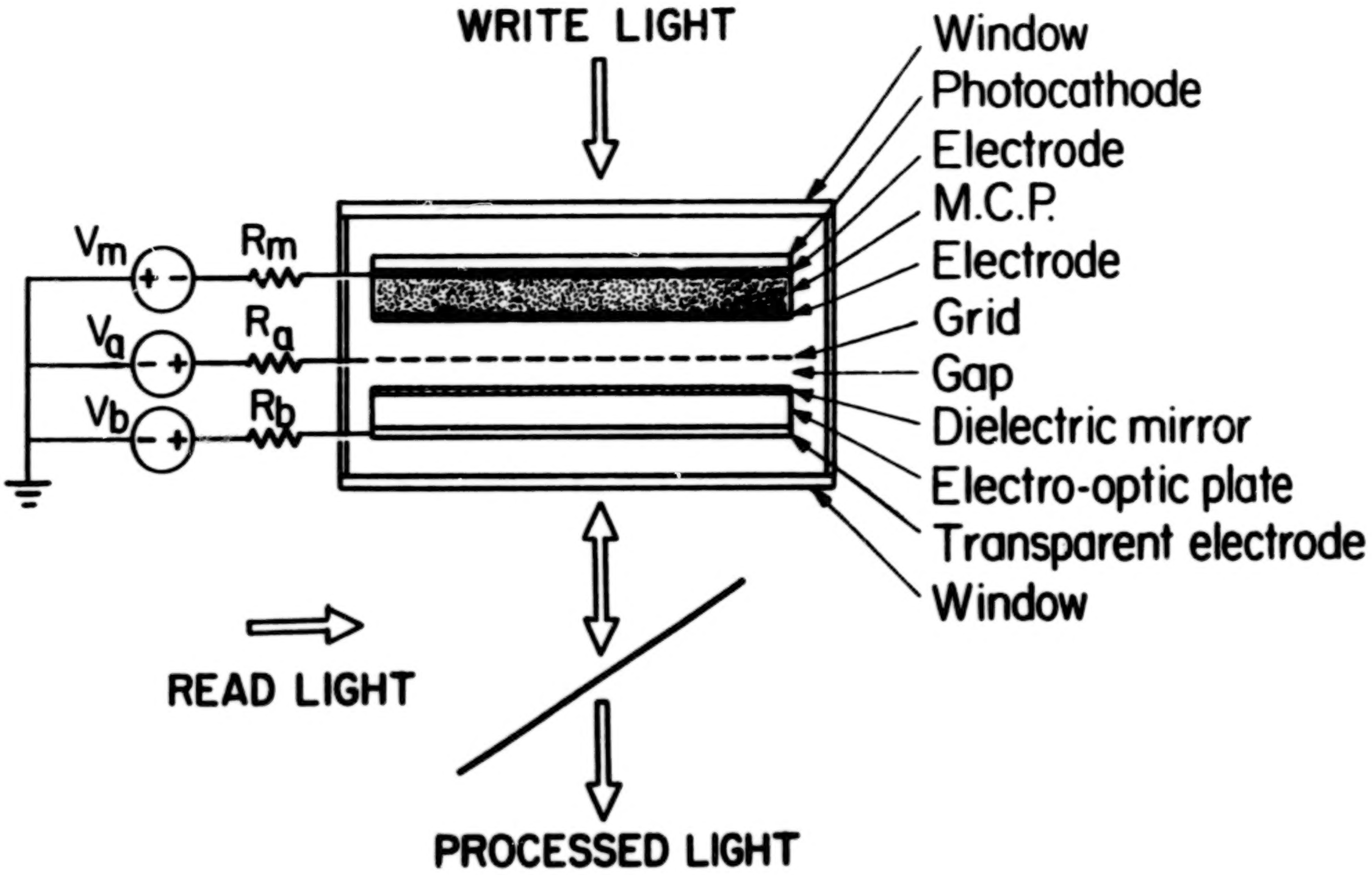

WRITE LIGHT
Window
Photocathode
Electrode
M.C.P.
Electrode
Grid
Gap
Dielectric mirror
Electro-optic plate
Transparent electrode
Window
V_m
R_m
V_a
R_a
V_b
R_b
READ LIGHT
PROCESSED LIGHT

Figure 1.- The microchannel spatial light modulator.

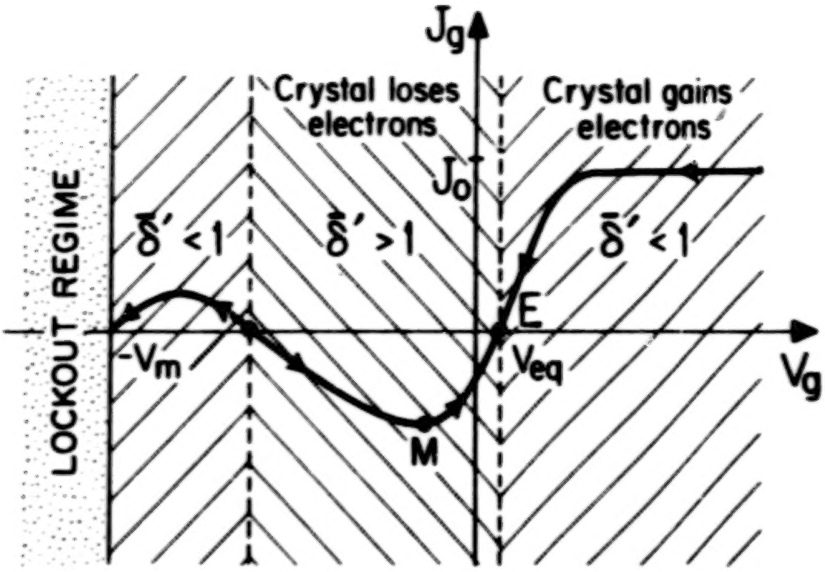


Figure 2.- Sketch of gap current density J_g in an MSLM as a function of gap voltage V_g for constant write light intensity I and MCP bias voltage $-V_m$.

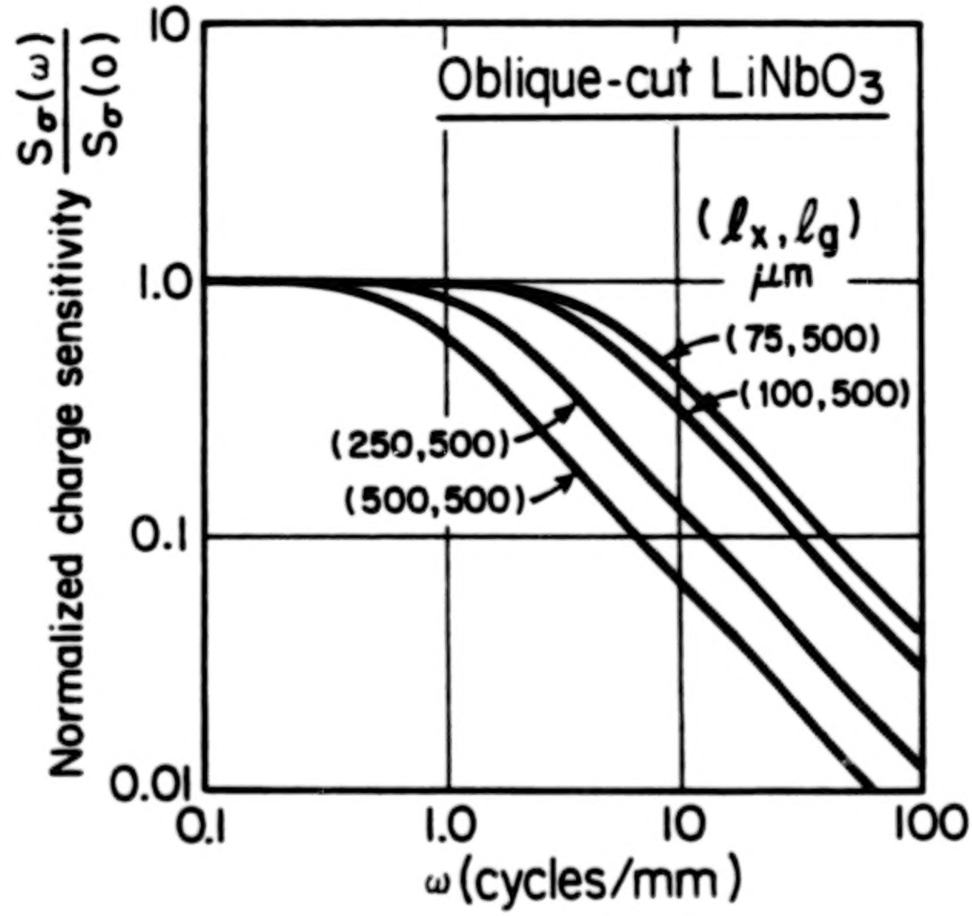
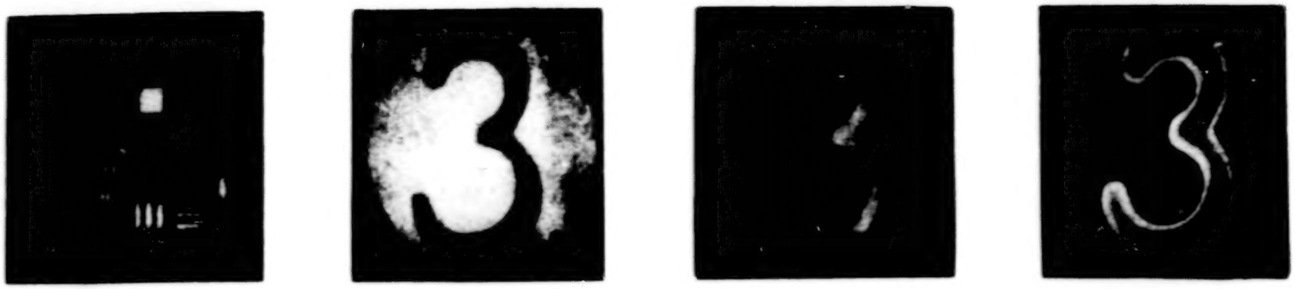


Figure 3.- Calculated normalized charge sensitivity $S_\sigma(\omega)/S_\sigma(0)$ plotted against spatial frequency ω for an oblique-cut $LiNbO_3$ MSLM with crystal thickness and gap width as parameters.



RESOLUTION CHART

CONTRAST REVERSAL

EDGE ENCHANCEMENT

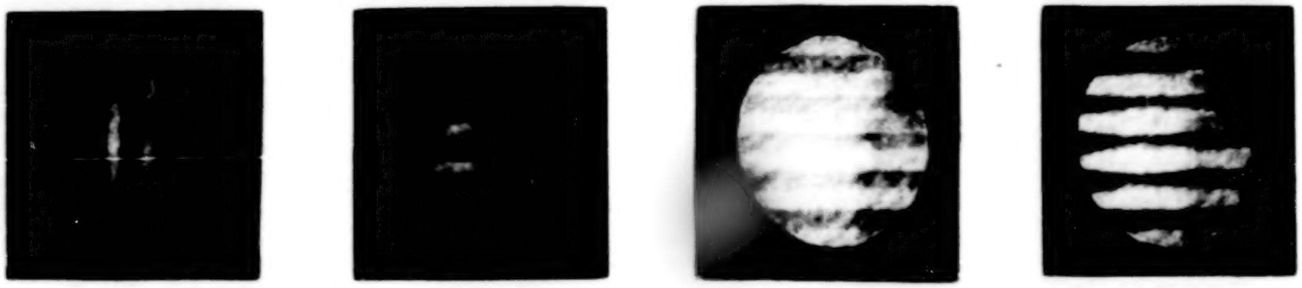


IMAGE A

IMAGE B

IMAGE B PLUS BACKGROUND

CONTRAST ENHANCEMENT

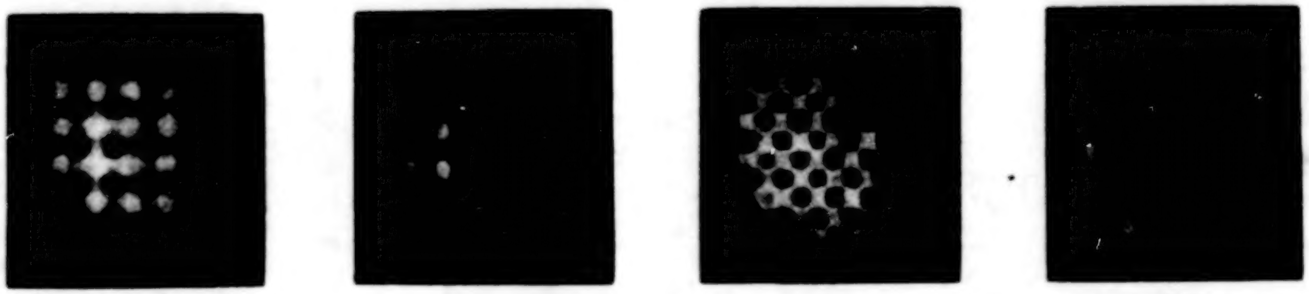


IMAGE ADDITION

AND

XOR

NOR

Figure 4.- Sample of some of the image processing operations achieved with the MSLM.

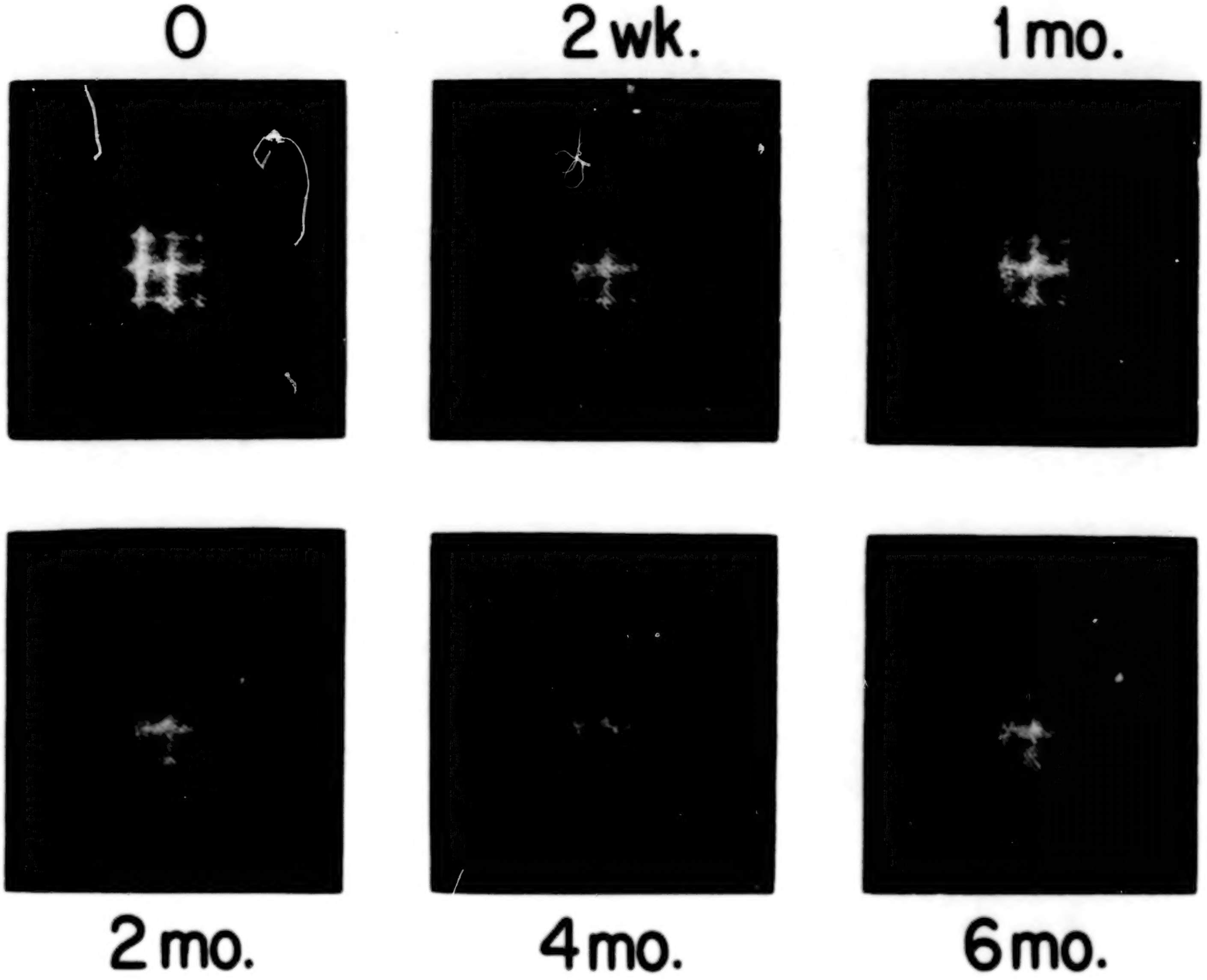

0
2 wk.
1 mo.
2 mo.
4 mo.
6 mo.

IMAGE DEGRADATION DURING LONG-TERM STORAGE

Figure 5.- Image degradation during long term storage in a vacuum-sealed visible-photocathode MSLM with a Z-cut $LiTaO_3$ crystal.

OPTICAL DATA LINKS USING LONG WAVELENGTH TECHNOLOGY*

D. J. Channin
RCA Laboratories
Princeton, NJ 08540

The replacement of short wavelength (0.82 to 0.86 μm) sources by long wavelength (1.1 to 1.6 μm) offers opportunities for data links with enhanced performance and improved reliability. Performance improvement comes from the reduced Rayleigh scattering at long wavelengths and the opportunity to set the source wavelength at the fiber dispersion minimum. Reliability comes from two factors: (1) the potential reduction in damage inherent in lower photon energies and (2) the opportunity to replace lasers by LED's for many high performance applications. This paper will focus on the latter approach.

This paper will give a review of RCA's InGaAsP/InP long wavelength (1.1 to 1.7 μm) laser diodes, LED's and PIN photodiodes. Performance characteristics of these devices will be presented with emphasis on their capability for data link service. At the present time, the use of edge-emitting LED's is favored for moderate link lengths such as are found in distributed computation and data base management systems.

Analysis of data link performance using present and projected components will be given. Figure 1 indicates the comparative performance of 0.84 μm lase diode/Si APD data links and 1.3 μm LED/InGaAsP PIN systems. Effects of dispersion are taken into consideration. The conclusion is that for service below 300 Mb/s, LED based fiber optic systems can better the performance of laser diode technology.

---

* Because this paper was not available at the time of publication, only the abstract is included.

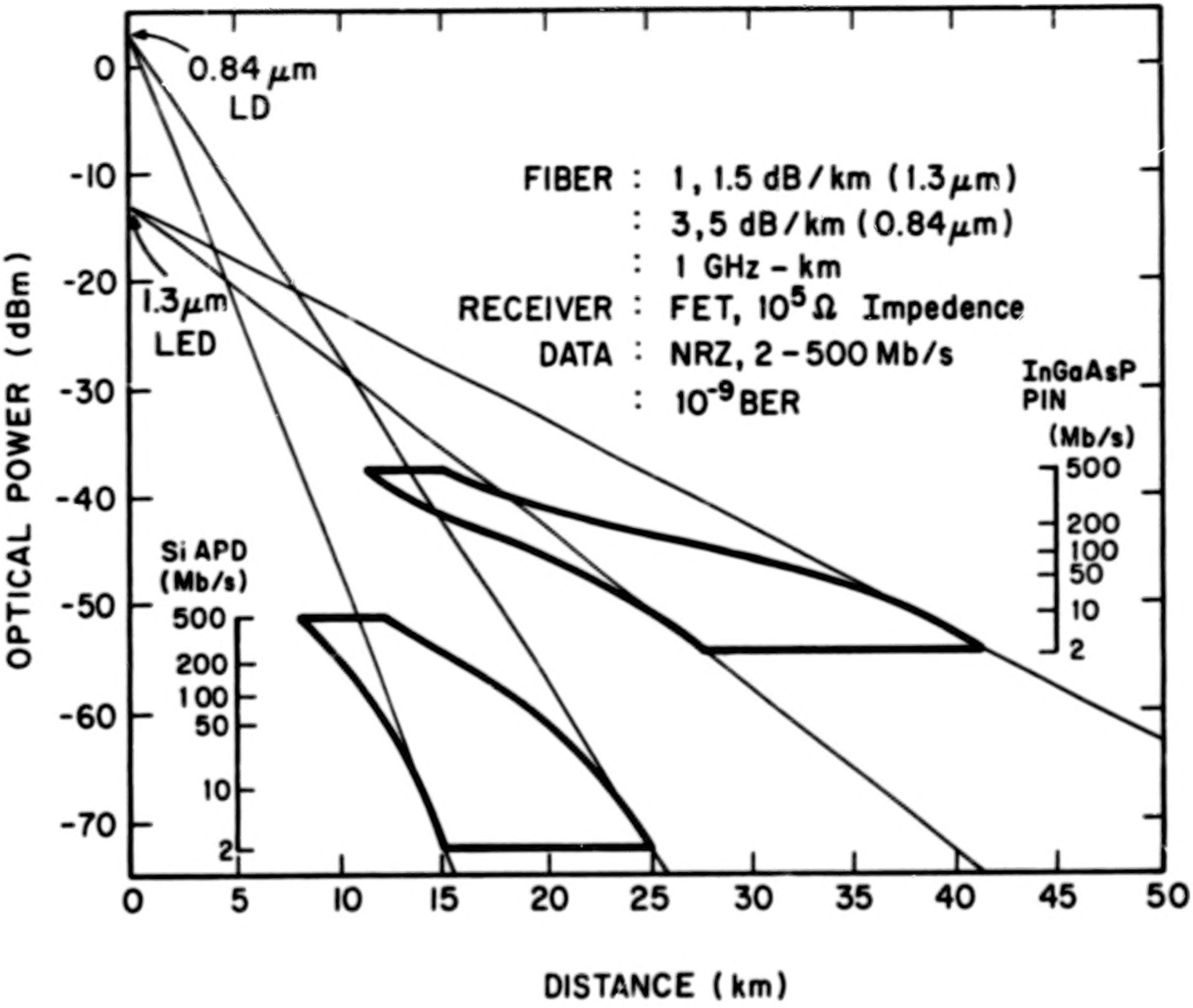


Figure 1.- Transmission characteristics of fiber optic data links.

INFRARED FIBER OPTIC FOCAL PLANE DISPERSERS

John H. Goebel
National Aeronautics and Space Administration
Ames Research Center

SUMMARY

Middle and far infrared focal plane detectors are currently less integrated in their design than are those in the near infrared. The difficulties of low temperature processing steps on the ultrahigh purity germanium detector material, which is required in the far infrared, limit the realization of integrated circuit technology in the near future. As an alternative approach, we discuss the utilization of far infrared transmissive fiber optics as a component in the design of integrated far infrared focal plane array utilization. A tightly packed bundle of fibers is placed at the focal plane, where an array of infrared detectors would normally reside, and then fanned out in two or three dimensions to individual detectors. Subsequently, the detectors are multiplexied by cryogenic electronics for relay of the data.

A second possible application is frequency upconversion ($v_1 + v_2 = v_3$), which takes advantage of the non-linear optical index of refraction of certain infrared transmissive materials in fiber form. Again, a fiber bundle is utilized as above, but now a laser of frequency v_1 is mixed with the incoming radiation of frequency v_2 within the non-linear fiber material. The sum, v_3, is then detected by near infrared or visible detectors which are more sensitive than those available at v_2. Due to the geometrical size limitations of detectors such as photomultipliers, the focal-plane dispersal technique is advantageous for imaging upconversion.

INTRODUCTION

Currently, realizations of instruments based upon infrared upconversion are limited by the availability of large high quality non-linear crystals. I describe here how fiber bundles of the non-linear material can substantially eliminate the availability problem. Such "pixelized" crystals, i.e., the fiber bundles, can considerably increase the conversion quantum efficiency available at low pump power densities. The integration of optical components thru the use of fiber optics can considerably add to the ruggedness and usefulness of instruments based on their incorporation when compared with those utilizing discrete components.

DISCUSSION

Linear Fiber Optics

Infrared focal planes are now in a phase of extreme integration pressure. Users are not satisfied with a few good detectors which scan a scene to thereby construct a picture. Rather the availability of microcircuit technology has created a large demand for CCD or CID focal plane arrays which "pixelize" a scene with considerable facility and accuracy. The cost investment to produce custom arrays is considerable to the point of being prohibitive in many instances. Thus only a few standard array formats have become available, and those are in silicon for the visible and HgCdTe for the infrared.

In certain imaging applications the detector pixel size is not matched conveniently to the spatial resolution of the telescope. Several factors affect the spacial resolution; most notably they are "seeing", diffraction, and optical quality. Any of these can severely degrade the resolution so that several detectors measure several spatial resolution elements or vice versa. This results in a loss of information.

In Figure 1 we show a proposal for coupling dissimilar spatial resolution scales so that in effect they are equivalent. A conical fiber bundle of appropriate length is matched at one end to the spatial resolution scale of the focal plane and at the other end to the detector array scale. Note that if the detector size is much smaller than the resolution size, care must be exercised in the cone design so as not to totally retroreflect the light in the cone before it exits onto the detectors.

The scheme outlined in Figure 1 can be used advantageously in the visible, near IR, and middle IR, where integrated circuit technology produces suitable detector arrays. In the far IR, the silicon and HgCdTe detectors do not function. There are, so far, only germanium based detectors which are physically sizeable as well as discrete. In this case a solution is also possible using fiber optics of appropriate materials and is shown in Figure 2. Suitable fiber materials are KRS5 and AgBr. The cone is tied at the focal plane end, while the detector end is free as in a broom construction. Individual detectors are then free to be placed over an appropriate spatial scale with due consideration to such ancillary components as integrating spheres, preamps, load resistors, and stressing clamp. It should be noted that fiber bundles are advantageous in space applications for routing focal plane images into compartments shielded from cosmic rays.

Nonlinear Fiber Optics

One of the really promising developments in optical materials in the last decade is the explosion of our knowledge about the properties and applications of non-linear optical crystals. Since non-linear properties were first investigated, it has been realized that one of their truly original applications is in the conversion of infrared radiation into visible where extremely sensitive detectors are available. There have been several successful instruments built for this purpose, but all suffer from one or another limitations imposed by the availability of pump sources or that of the requisite non-linear material. It appears that some of these major problems can be alleviated through the introduction of non-linear fiber bundles as proposed here.

First I'll list the basic theoretical description of non-linear upconversion. If photons of two different frequencies, v_1 and v_2 are simultaneously propogating thru a uniaxial crystal with an ordinary index of refraction n^o and an extraordinary index n^e, then the crystal can be rotated to a specific angle Θ_m, the mixing angle, such that the laws of conservation of energy and momentum for the waves can be simply related as

$$v_1 + v_2 = v_3 \quad \text{energy conservation}$$

and

$$\vec{k}_1 + \vec{k}_2 = \vec{k}_3 \quad \text{momentum conservation}$$

i.e., a third wave at v_3 is produced. Polarization must also be considered and this leads to two types of momentum conservation or phase matching.

For Type 1 mixing both waves have the same ordinary polarization

$$n^{\circ}{}_1 v_1 + n^{\circ}{}_2 v_2 = n^{e}{}_3 v_3 \quad \text{Type 1}$$

For Type 2 phasematching, orthogonal polarizations are mixed

$$n^{e}{}_1\, v_1 + n^{\circ}{}_2 v_2 = n^{e}{}_3 v_3$$

For uniaxial crystals, the polarization can be expanded in terms of the electric field strength E

$$P_i = 2 \sum_{i,j,k} d_{ijk} E_j E_k$$

The non-linear interaction is proportional to the product of field of strength of the two waves and the crystal's effectiveness as a non-linear medium is measured by d_{ijk}. Under certain assumptions about the symmetry of the polarization matrix, e.g., Kleinman symmetry, the experimental equivalent of d_{ijk} is d_{eff}. In practice it is almost impossible to evaluate d_{ijk} theoretically.

For an instrument, the practical parameter of interest is the quantum conversion efficiency η

$$\eta = \frac{512\ \pi^5\ d_{eff}\ I_L l^2}{n_1 n_2 n_3\ \lambda_2 \lambda_3\ C} \left[\frac{\sin x}{x}\right]^2$$

where $x = \frac{(\Delta k l)}{2}$ and $\Delta k = \left| \vec{k_3} - \vec{k_1} - \vec{k_2} \right|$

I_L is the laser pump intensity usually chosen at v_2 and l is the interaction length of the crystal. The term in brackets is related to the coherence length of the interaction X which is dimensionless. Of immediate practical interest are the important Manley-Rowe relationships which govern the power flow from wave to wave

$$\frac{\Delta P_1}{v_1} = \frac{\Delta P_2}{v_2} = \frac{\Delta P_3}{v_3}$$

which essentially state that 100% power conversion of infrared into visible photons is possible. Indeed that has been shown to be possible in the laboratory.

Instruments have been demonstrated based upon different non-linear crystals to produce infrared upconversion with respectable NEP's and η's. Boyd and Townes (1977) have used proustite ($Ag_3As\ S_3$) to make an astronomical imager with an NEP = 3×10^{-10} W/$\sqrt{Hz}$ and an $\eta = 2 \times 10^{-7}$. Images were obtained of the Sun, Moon, Mercury and VY CMa. Voronin et al (1975) used $AgGaS_2$ to achieve 40%-η with a high purity (α (1.06 μm) < 0.1 cm^{-1}) crystal able to utilize high laser energy levels. Note that η depends only on the pump power, not the signal source power. Hence the great interest in upconversion for low level signal detection. Gurski (1973) used a $LiIO_3$ crystal to convert 3.4 μm radiation with 100% efficiency. He also achieved an NEP = 1×10^{-15} W/$\sqrt{Hz}$. Both Voronin and Boyd were able to achieve imaging, as have other workers.

The requirements of high field strengths from the pump necessary to drive the non-linear interaction requires focusing the pump and signal source in the crystal to the threshold of damage, usualy at 10's of MW/cm^2 in presently available materials. Such high energies are usually available only from pulsed lasers of low duty cycle. This situation is a disadvantage in low level signal detection and leads to the apparently low n and NEP of Boyd and Townes.

From the expression for η, the easily controlled parameters of an upconverter are d_{eff}, I_L, and l. While η grows linearly with I_L, it grows as the square of d_{eff} and l. The search for materials with higher d_{eff} continues, but dramatic order of magnitude leaps are very unlikely. Present values of d_{eff} are limited to about 300×10^{-9} esu. Specific application may limit the useful d_{eff} to values less than 100×10^{-9} esu.

Refinement of material quality is more easily achieved than inventing new materials. Present absorption coefficients are usually greater than 0.1 cm^{-1}, limiting practical crystal usage to 10 cm lengths. However, because of the need to focus in crystals to up the field strength, the useful portion of crystal is in practice much less than this figure. The spectral bandpass Δv_{bp} of an upconverter is determined by the crystal length and is approximately $\Delta v_{bp} = l^{-1}$ in cm^{-1}.

In Figure 3, I propose that fiber bundles of non-linear material can be used to alleviate some of the above restrictions on the performance of infrared upconverters. The fibers are able to maintain a given pump intensity I_L over greater lengths in a focused beam profile. The full length of a fiber can be employed usefully to increase n. If in the figure the collimated beams of v_1 and v_2 are focused on the crystal fiber bundle, then the image is preserved throughout the detection system. Longer lengths of fibers employed mean increased spectral resolution. Fiber bundles can also be formed from selected high quality lengths of fibers.

At the moment, the most promising material available for production of upconversion in fibers is AgGaS$_2$. Most all of its properties are known which are essential to production of such fibers. In a collaborative endeavor with R. Byer and R. Feigelson of Stanford University, D. Dimiduk, and myself, with the support of the Ames Research Center's Directors Discretionary Fund, are working to the realization of a fiber optical non-linear upconverter, with the goal in mind of achieving single photon counting in the 5-m region of the infrared with imaging capability.

REFERENCES

Boyd, R. W. and Townes, C.H,. Applied Physics Letters 31, 440 (1977)

Gurski, T. R., Applied Physics Letters 23, 273 (1973)

Voronin, E.S. et al, Soviet Journal of Quantum Electronics 5, 597 (1975)

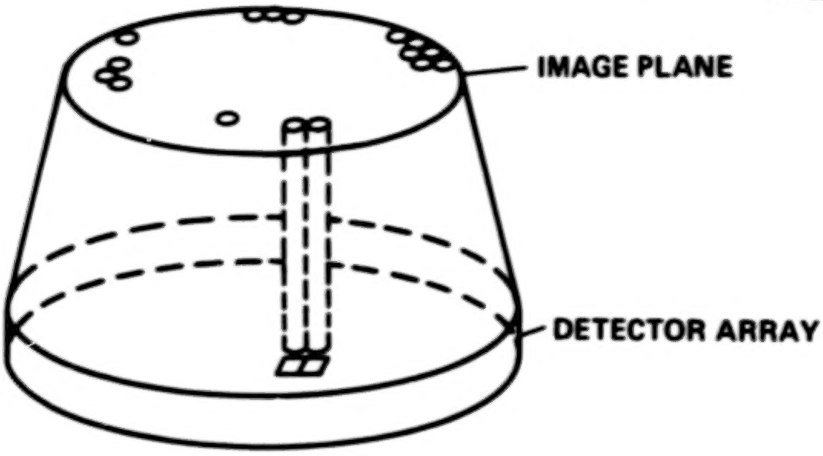


IMAGE PLANE
DETECTOR ARRAY

SUITABLE FOR
- **GLASS FORMING MATERIALS**

Figure 1.- Mid near IR imaging focal plane disperser.

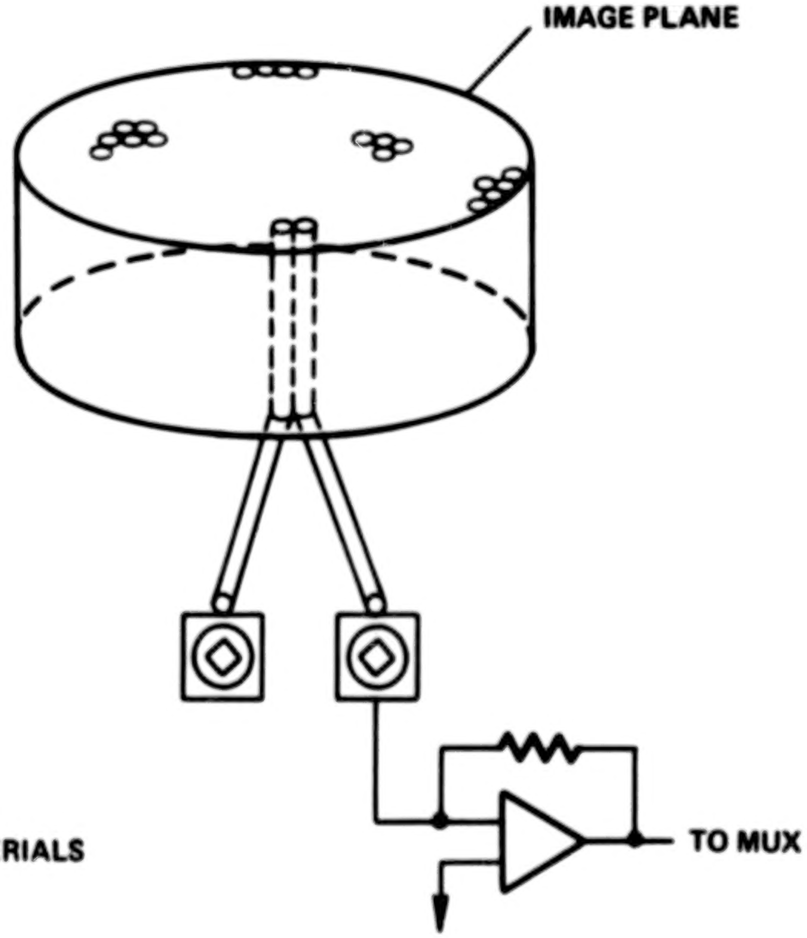


IMAGE PLANE
TO MUX

SUITABLE FOR
- **POLYMER MATERIALS**
- **CRYTASTALLINE MATERIALS**

Figure 2.- Far IR imaging fiber optic focal plane disperser.

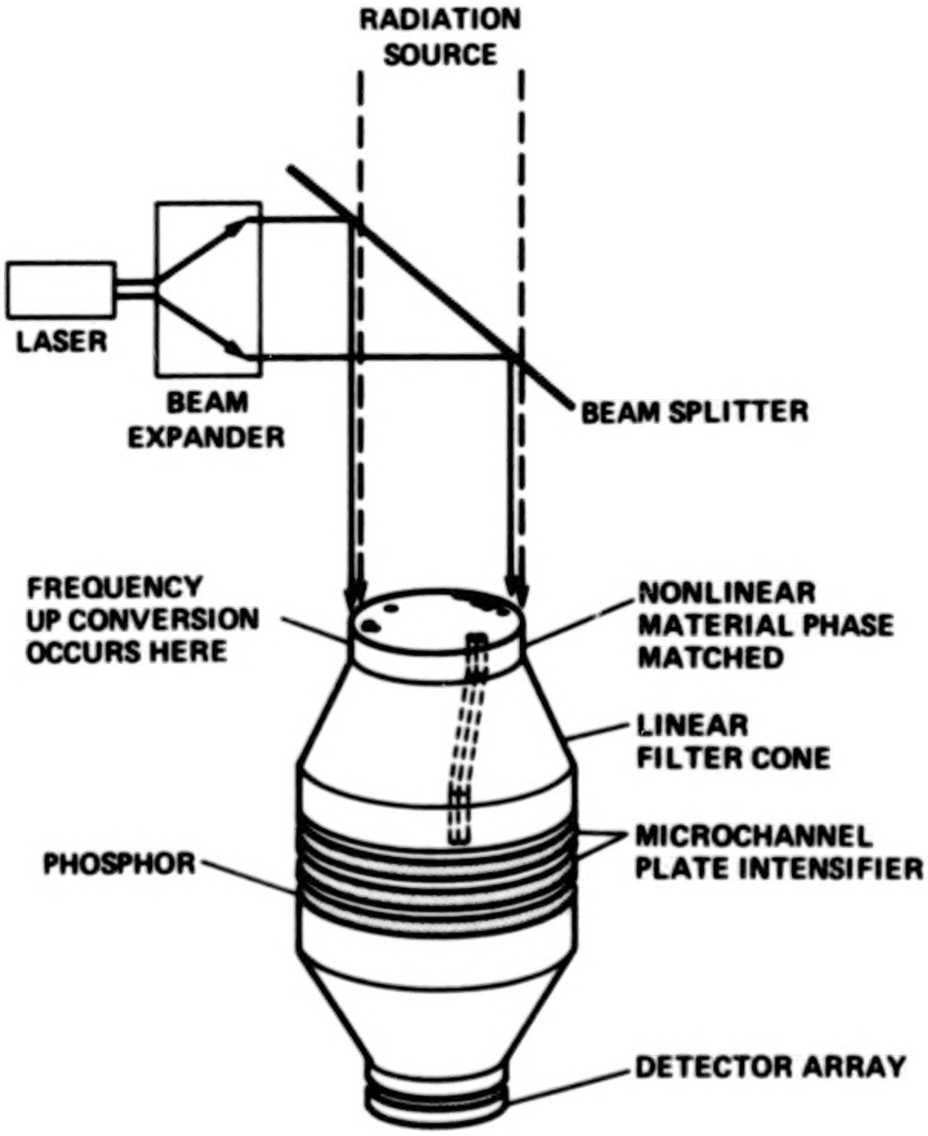

RADIATION SOURCE
LASER
BEAM EXPANDER
BEAM SPLITTER
FREQUENCY UP CONVERSION OCCURS HERE
NONLINEAR MATERIAL PHASE MATCHED
LINEAR FILTER CONE
MICROCHANNEL PLATE INTENSIFIER
PHOSPHOR
DETECTOR ARRAY

Figure 3.- Mid IR focal plane disperser used for imaging frequency upconversion.

E-BEAM GENERATED HOLOGRAPHIC MASKS FOR OPTICAL VECTOR-MATRIX MULTIPLICATION+

Steven M. Arnold
Honeywell Corporate Technology Center
Bloomington, Minnesota 55420

Steven K. Case
Department of Electrical Engineering
University of Minnesota
Minneapolis, Minnesota 55455

SUMMARY

A proposed optical vector-matrix multiplication scheme encodes the matrix elements as a holographic mask consisting of linear diffraction gratings. The binary, chrome-on-glass masks are fabricated by e-beam lithography. This approach results in a fairly simple optical system that promises both large numerical range and high accuracy. A simple holographic mask has been fabricated and tested.

INTRODUCTION

There has recently been considerable interest in optical computing since it offers very high computation throughput rates for mathematical operations amenable to parallel computation. One class of such operations, vector-matrix multiplication, can be used for performing discrete Fourier transforms, coordinate transformations, pattern classification, and many other computations. The general vector-matrix multiplication may be written as:

$$y_m = \sum_{n=1}^{N} H_{mn} x_n \quad (m = 1,2,\ldots M)$$

One optical approach to performing this computation uses N light sources to represent the components x_n of the input vector, M detectors to represent the components y_m of the output vector, and suitable optics to assure that a fraction H_{mn} of the light from source x_n gets to each detector y_m. The problem can be suitably scaled so that all parameters fall within acceptable positive ranges. Optics to perform the function of the matrix H generally will be fixed, while the sources are modulated to represent various input vectors x.

In principle, the performance of this optical computer is dependent on a number of considerations involving the optics, detectors and sources. In practice, numerical range and accuracy will usually be limited by matrix element imperfections, while speed will be limited by the amount of light reaching the detectors. For this reason, our work has focused on efficient optics to precisely distribute light among the various detectors.

In most schemes for optical vector-matrix multiplication, the matrix is encoded as a rectangular array of apertures or gray tones in a mask. This approach, represented in Figure 1, encounters several limitations. A complex optical system is required in order to illuminate and receive light from specific columns and rows of the matrix mask. Much light is discarded in

---

+This work was sponsored by the Air Force Office of Scientific Research under Contract No. F49620-80-C-0029.

providing uniform illumination to the mask, with the mask passing only about half of what remains. Numerical range and accuracy are limited by the space-bandwidth product of the mask (generally less than 10^6 with conventional plotting techniques). Small matrix elements result in small apertures with low relative accuracies. If results differ from those intended, it is generally difficult to modify a mask except by starting anew.

PCGH CONFIGURATION

Our approach to optical vector-matrix multiplication is based upon an e-beam generated diffractive mask which we call a partitioned computer generated hologram (PCGH) (see Figure 2). Each of N PCGHs is illuminated by collimated light from a single element of the source array and thus represents one column of the NxM matrix mask depicted in Figure 1. Each PCGH is partitioned into M linear gratings which diffract light to the M detectors. The optical power diffracted by a particular grating is made proportional to the value of the required matrix element.

Figure 3 illustrates a PCGH intended to produce 10 equal intensity outputs when uniformly illuminated. This PCGH contains 10 equal area gratings, each with its own spatial frequency. Facets are arranged symmetrically about the center to provide immunity to beam wander.

The PCGHs are fabricated as binary chrome-on-glass holograms where the pattern is delineated via e-beam lithography. A glass plate is first coated with a layer of chrome and a layer of e-beam resist. A pattern is exposed in the resist by e-beam direct writing and the resist is developed. The developed resist then serves as a mask for etching the pattern into the chrome.

Our e-beam PCGH optical vector-matrix multipliction scheme has several advantages over the scheme in Figure 1. E-beam lithography offers a higher space-bandwidth product, which can translate into greater numerical accuracy. Also, the PCGHs are in the Fourier plane of the transform lens with respect to the detectors. This means that the only requirement for light to reach a particular detector is that it be traveling in the right direction upon leaving the PCGH. Therefore, the input modules, consisting of source, collimating lens and PCGH, may be located anywhere within the aperture of the transform lens. This same immunity to shifts allows a PCGH to be partitioned into facets in any manner consistent with dividing up the available light amongst the various detectors (providing, of course, that the facets do not become too small). Other advantages relate to optical efficiency. All light striking the PCGH can be used. Light need not be wasted in achieving uniform illumination; non-uniform illumination is acceptable so long as its effects are accounted for in the partitioning. Small facets associated with lesser outputs can be made physically larger by placing them where PCGH illumination is lowest. The various considerations which go into the design of a PCGH are discussed in the following sections.

PCGH DIFFRACTION ANALYSIS

Each facet of the PCGH contains a linear grating to diffract incident light to the appropriate detector. The spatial frequencies of these linear gratings are determined by the system geometry. First-order diffracted light from a facet of spatial frequency ν will be focused in the detector plane a distance $\nu\lambda F$ from the transform lens axis (Figure 1), where λ is the wavelength and F the transform lens focal length. Our design is for a 10-element linear detector array. This requires 10 equally spaced grating frequencies. The widest possible detector separation is achieved for grating frequencies $n\Delta\nu$, where $\Delta\nu$ is the frequency separation and $n = 10, 11, 12, 13, 14, 15, 16, 17, 18,$ and 19. Then the unwanted harmonic frequencies from the square wave gratings begin at $20\Delta\nu$ and will not coincide with the desired outputs.

MODULAR FACETS

The matrix values are encoded into the PCGH via grating area modulation. The hologram must therefore be divided into facets such that the amount of light diffracted by a facet to its detector is proportional to the required matrix element. Various considerations lead us to partition the PCGH into facets along a square grid.

Mathematically, the transmittance of a facet can be regarded as the product of its aperture and an infinite linear grating. By the Fourier convolution theory, the diffraction pattern of this facet is the diffraction pattern of its aperture convolved with the delta function from the infinite linear grating. In other words, the effect of the linear grating is to shift the location of the diffraction pattern of the facet aperture.

Figure 4 indicates the diffraction pattern due to a square aperture of dimension D. The main lobe has a width $2\lambda F/D$ and contains 81.5% of the energy passing through the aperture, which we label as 0 dB. The sidelobes form a rectangular array with sidelobe energy diminishing inversely as the square of the distance from either axis. The figure indicates the energies (in dB) of the various sidelobes relative to the main lobe. It is seen that their energies diminish most rapidly along a diagonal.

Crosstalk between channels depends on how these diffraction patterns overlap adjacent detectors. As indicated earlier, the separation of detectors is $\Delta\nu\lambda F$. As facets are made smaller, their diffraction patterns become larger, requiring higher grating frequencies to separate them. For this reason, we impose a minimum square facet size, based on detector separation and an acceptable level of crosstalk. We size our detector aperture to capture only the main lobe of this minimum facet diffraction pattern. A larger aperture would capture lesser lobes of the channel of interest, but also some greater lobes of adjacent channels, thereby degrading the signal-to-noise ratio. Also, zero intensity at the edges of the detector aperture eases mechanical tolerances.

In the diffraction pattern of Figure 4 the square box in the center represents the detector aperture. Similar boxes are used to indicate possible locations of adjacent detectors. These have been separated diagonally to use the more rapid sidelobe decay (implying grating fringes which run diagonally within the facets). A separation of $2\lambda F/D$ in each dimension results in 29 dB of crosstalk with some very bright axial sidelobes just outside the detector aperture. A separation of $3\lambda F/D$ yields a much more comfortable 37 dB crosstalk. With this as our choice, the minimum facet size will be $D = 3\sqrt{2}/\Delta\nu$.

The above discussion of crosstalk implies that each facet is a minimum square facet. This would be a severe constraint on system numerical range. In practice, we form the facet for each channel from many minimum sized modular subfacets. Therefore, the diffraction pattern for a single channel is not that of a single subfacet, but rather the pattern is due to the aperture consisting of all subfacets for that channel. Crosstalk can be minimized by clustering all required subfacets of a given spatial frequency into one or two large facets with minimum perimeter, such as was done in the PCGH of Figure 3.

To have the greatest flexibility in partitioning the PCGH, we would like the modular subfacets to be as small and numerous as possible. Since they can be no smaller than $D = 3\sqrt{2}/\Delta\nu$ (with our 37 dB crosstalk limit), we want a large spatial frequency separation $\Delta\nu$. However, a large spatial frequency separation implies large spatial frequencies and hence small grating periods. If the grating period becomes comparable in size to the e-beam spot size, considerable grating duty cycle errors with corresponding diffraction efficiency errors will result. For these reasons we have elected to start with grating frequencies of 60 to 114 lp/mm (measured along either axis) and 0.5 mm subfacets. This gives us 400 subfacets in a 1 cm x 1 cm PCGH and a maximum numerical accuracy of about 20 dB. With Gaussian illumination, the effective number of subfacets is extended to more than 5000 since the corner facets have about 6% the intensity of illumination of the central facets.

TRIMMING METHODS

To extend our numerical accuracy beyond about 20 dB, it is necessary to employ a separate lithography step to adjust the relative amplitudes of the outputs. The ability to do this trimming is one of the chief advantages of e-beam lithography and the PCGH. There are several possible methods of trimming a PCGH once it has been made and tested. The best trimming method is to add a negative facet, that is to say, a facet exactly out of phase with the existing facet. This is much easier than it might sound: the negative facet can be written in space already occupied by the existing positive facet, leaving a completely open area. Linewidth and phase problems are overcome since, if we write over the whole area, the phase and duty cycle will be exactly the complement of what was already there. The easiest and probably best way to trim, because it minimizes registration requirements, is to draw a square box of the appropriate size somewhere near the middle of the existing facet.

PARTITIONING

The task of partitioning the PCGH to achieve the correct relative outputs is greatly simplified by the decision to adopt module facets. The intensities of rectangular facets are easy to compute for either uniform or Gaussian illumination. We need only determine which subfacets are to be assigned to which channels. We do this by means of an algorithm which maximizes accuracy and minimizes crosstalk. Each subfacet is assigned entirely to one channel. Trimming is then used to achieve the desired numerical accuracy. With a 1 cm x 1 cm PCGH, 0.5 mm subfacets and a Gaussian illumination diminishing to e^{-2} at the edges, a subfacet in the corner will receive only 0.00019 of the total energy. Therefore, it is possible to obtain as much as 37 dB of numerical range, with no channel receiving less than a full subfacet.

EXPERIMENTAL RESULTS AND DISCUSSION

The experimental setup for demonstation and evaluation of PCGHs is indicated in Figure 5. Light from a He-Ne laser is passed through a pinhole spatial filter to create a point source which is then imaged in the detector plane by a lens. The PCGH is kinematically mounted in a micropositioner placed immediately after this lens. This mount is designed to allow removal for trimming and subsequent replacement of the PCGH without disturbing the alignment. The distance from the PCGH to the detector plane is 50 cm. In the detector plane, we have a single UDT-455 photodetector mounted on a motorized translation stage. One of several square apertures, selectable in size from 1 to 2 mm, is placed in front of this detector to define and limit its effective area. A second identical photodiode directly monitors the laser output. By ratioing the signals of these two detectors, we eliminate a number of potential noise sources, including laser power fluctuations. The two detector signals can be fed into a logarithmic amplifier for strip chart traces and rough measurements, or else the signals can be measured directly when the greatest accuracy is required.

Our experimental work thus far has been aimed chiefly at constructing experimental apparatus capable of yielding the precise measurements needed for verifying the 40 dB numerical accuracy and/or 40 dB numerical range we seek. In this regard, we consider the setup as described to be adequate with only minor refinements.

We have fabricated the PCGH of Figure 3. This hologram utilizes diagonally separated outputs, compact facets and symmetry about the axis. Figure 6 shows the resultant diffraction pattern at the output plane. Since the grating fringes are written at 45° relative to the facet boundaries, the many on-axis sidelobes from each output are seen to be diagonally separated from the adjacent output signals. Figure 7 is an intensity trace through the centers of the outputs. The spacing of the outputs was selected to give 37 dB worst case crosstalk between adjacent channels.

In light of our preliminary measurements, the PCGH method of optical computing appears capable of high numerical accuracy over a wide numerical range.

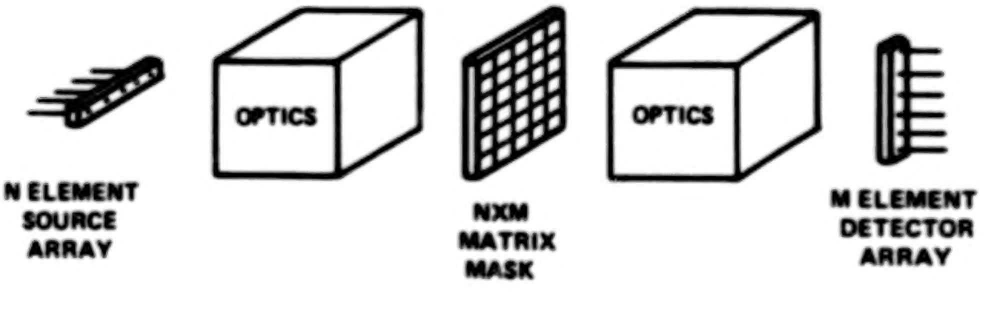

N ELEMENT
SOURCE
ARRAY
OPTICS
NXM
MATRIX
MASK
OPTICS
M ELEMENT
DETECTOR
ARRAY

Figure 1.- A general scheme for optical vector-matrix multipliation.

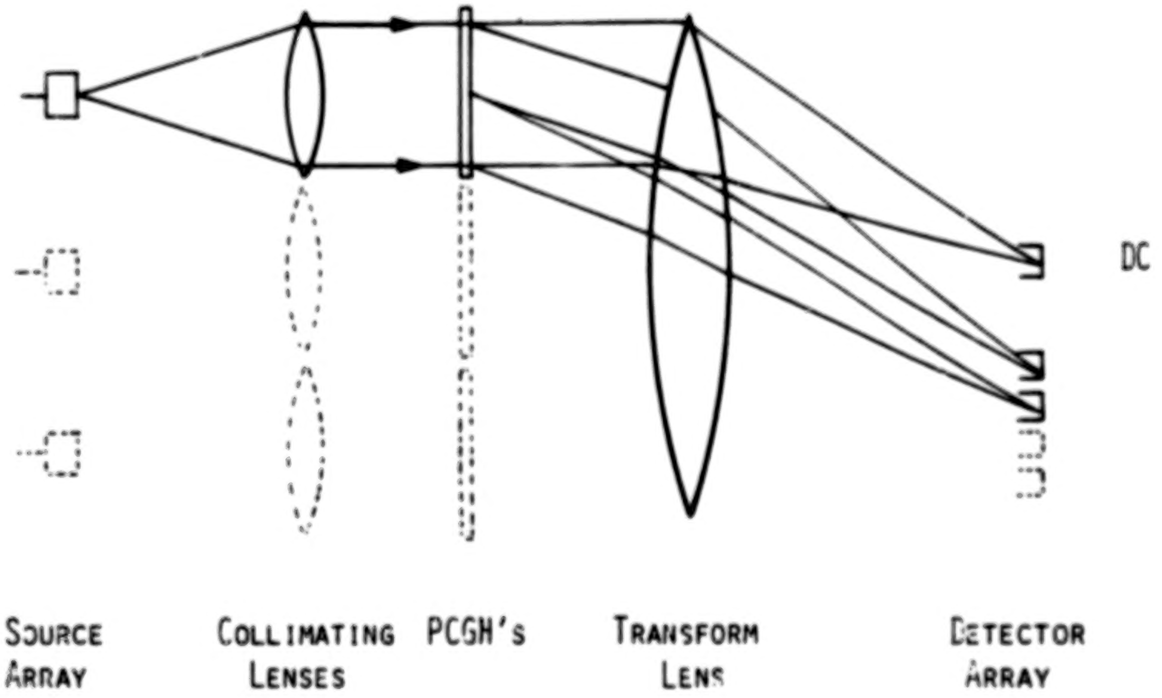

DC
SOURCE
ARRAY
COLLIMATING
LENSES
PCGH's
TRANSFORM
LENS
DETECTOR
ARRAY

Figure 2.- PCGH vector-matrix multiplier.

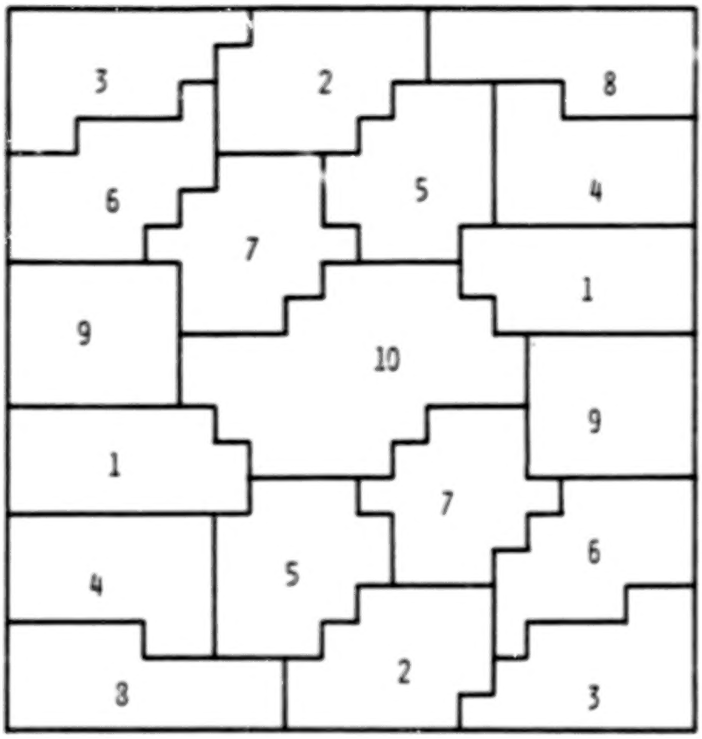

3
2
8
6
5
4
7
1
9
10
9
1
7
4
5
6
8
2
3

Figure 3.- PCGH design showing facet boundaries. (Fringes within facets not shown.)

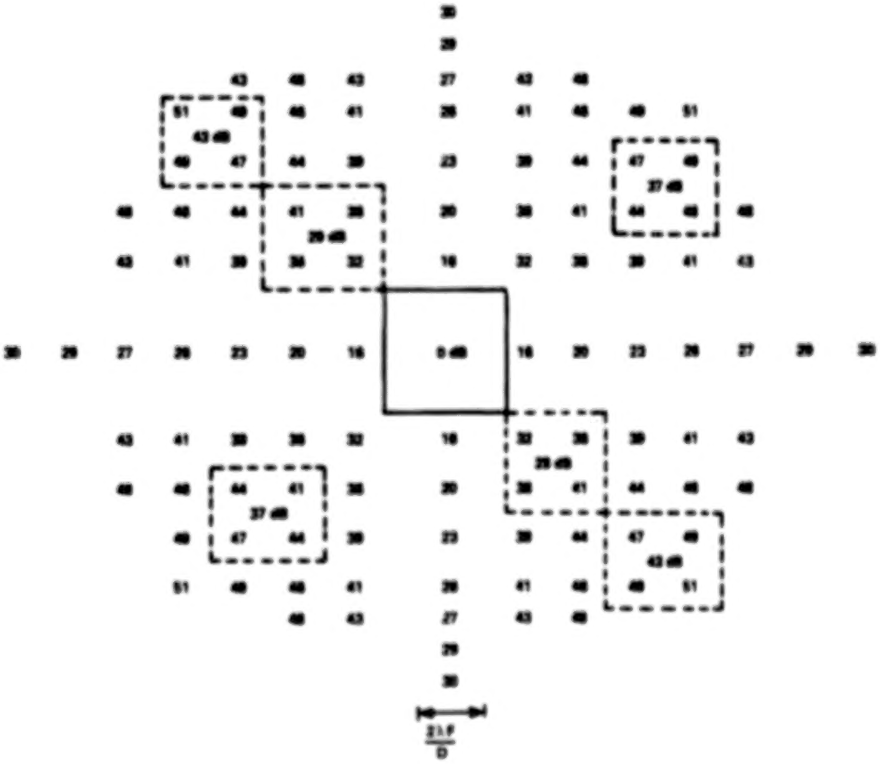

43 dB
37 dB
29 dB
0 dB
29 dB
37 dB
43 dB
2λF/D

Figure 4.- Crosstalk from square subfacets.

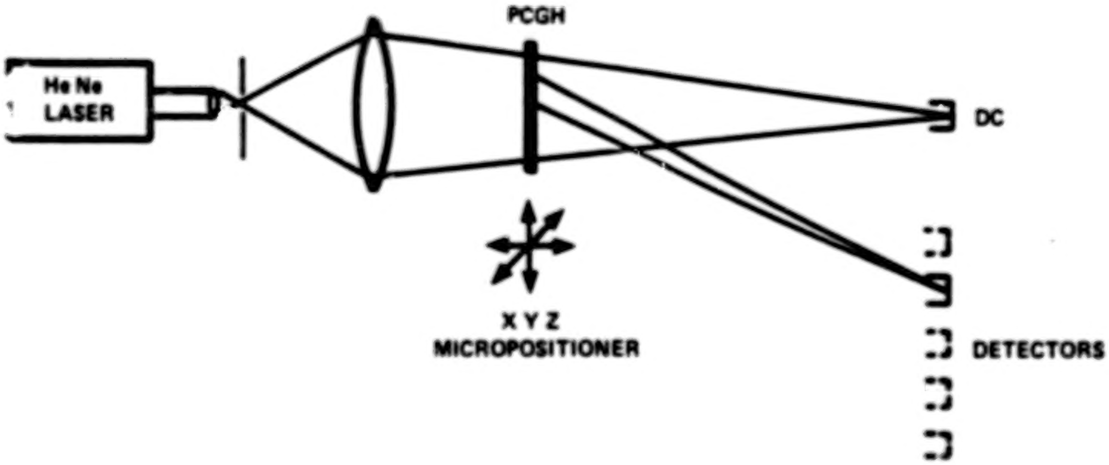

PCGH
He Ne
LASER
DC
X Y Z
MICROPOSITIONER
DETECTORS

Figure 5.- Experimental setup for demonstration and evaluation of PCGHs.

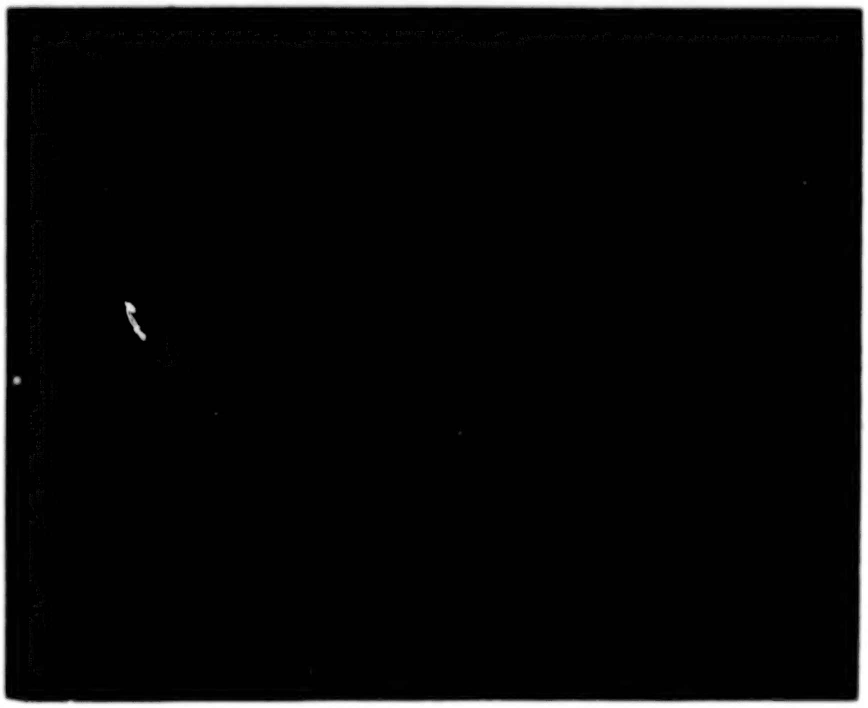


Figure 6.- Diffraction pattern from the PCGH. The undiffracted beam, all 10 first-order diffracted beams, and the first few second-order beams are visible.

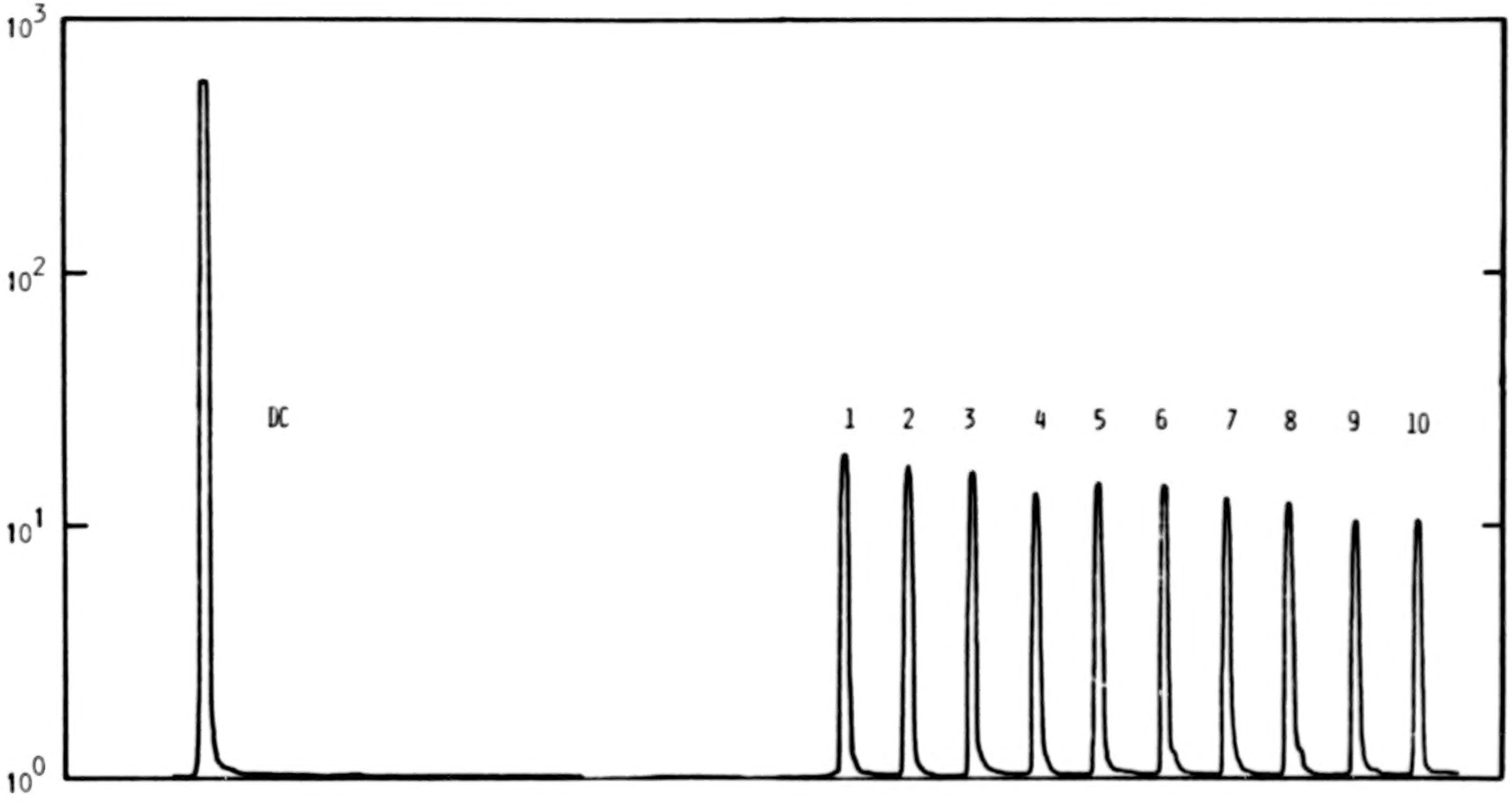


Figure 7.- Scan through diffraction pattern from the PCGH.

MULTI-FACET HOLOGRAPHIC OPTICAL ELEMENTS[†]

Steven K. Case and Paul R. Haugen
Electrical Engineering Department
University of Minnesota
Minneapolis, Minnesota 55455

SUMMARY

New types of holographic optical elements are demonstrated which combine some of the flexibility of computer generated holograms with the large space-bandwidth product and high diffraction efficiency of interferometrically recorded volume phase holograms. The optical elements are recorded by subdividing a volume hologram film surface into numerous small areas (facets), each of which is individually exposed under computer control. Each facet is used to produce a portion of the desired final wavefront. Three different optical elements are demonstrated.

INTRODUCTION

Holographic optical elements are an attractive alternative to conventional optical elements in systems using coherent light. The attributes of holographic optical elements include the capability of producing low f-number diffraction limited optics, and the ability to generate arbitrary wavefronts. They are also potentially easy to replicate and have low weight and compact size. Unfortunately, not all of these attributes are shared by all hologram types.

We could categorize holograms as computer generated, interferometrically generated, or generated by a hybrid technique as described in this paper. Computer generated holograms (CGH) offer the greatest flexibility in the production of arbitrary wavefronts. CGH's, however, usually have limited space-bandwidth products and diffraction efficiency. Volume phase holograms, on the other hand, can have nearly 100% diffraction efficiency into a single order and very large space-bandwidth products (necessary for low f-number optical elements), but offer limited flexibility in the production of arbitrary wavefronts.

The new type of hologram described here combines the best features of the previous hologram types: high efficiency, large space-bandwidth product, and greater flexibility in wavefront generation. The principle used to design the new hologram is straightforward. While the interference pattern required to record an entire "arbitrary wavefront" volume hologram is quite complex, the required pattern within a small area on the hologram is relatively simple such that it can be approximated by planar, spherical, or cylindrical waves. Thus our new holographic optical element is recorded by individually exposing numerous, small volume hologram facets. The wavefronts used to record each facet can be different but are produced by conventional optical elements which are positioned via computer control.

[†] This work supported by the University of Minnesota and by AFOSR/ARO Grant DAA29-81-K-0033.

WAVEFRONT TRANSFORMATION OPTICAL ELEMENTS

The first type of optical elements that we describe are used for wavefront transformations. For this system we start with a Gaussian intensity profile laser beam which we wish to transform into another shape in order to optimally illuminate some object (for interferometric testing or materials processing). In this example, we assume that we must illuminate the periphery of a hollow, square box. Such a transformation requires two holographic optical elements as shown in Fig. 1. The first hologram acts to spatially redistribute the input light beam and the second hologram produces a desired phase front on the redistributed light so that it can propagate toward a distant object.

The first hologram is recorded by subdividing the hologram surface into numerous, small squares (facets) each of which is individually exposed as in Fig. 2. The spatial frequency of the grating recorded within each facet is varied by suitably adjusting the object beam incidence angle before recording. The grating frequency within each facet is chosen such that the input beam in Fig. 1 will be suitably deflected to be spatially rearranged when it reaches the plane of Hologram #2. Design considerations for the hologram are given in Ref. 1 and will not be repeated here.

The redistributed light from Hologram #1 and a plane wave are used as the recording waves for Hologram #2. In this manner, Hologram #2 can rediffract the light arriving from Hologram #1 to produce a final output wave with the desired intensity distribution and a planar wavefront. Because both holograms are recorded in dichromated gelatin, they have nearly 100% diffraction efficiency so that the redistribution process is light efficient. In Fig. 3 we have photographed the light pattern in five equally spaced planes to show the gradual redistribution of light via our process.

SPACE-VARIANT LENS

The multi-facet hologram technique is also useful for the production of customized holographic lenses with very large space-bandwidth products. As a demonstration, we constructed a space-variant cylindrical lens capable of producing a line focus on a spherical surface.[2] The construction of a conventional lens that would perform this task would be extremely difficult if not impossible. Without multi-facet hologram approach, however, we divide the hologram film into a set of horizontal zones (facets) and specify the focal length that a cylindrical lens wihin each facet would need in order to bring a plane wave to a line focus on the object surface (see Fig. 4).

For recording, a thin, stationary, flexible mask containing a horizontal slit aperture is again pressed lightly against the film. Exposure is with a planar reference wave incident at approxiately 35° with respect to the film normal and a cylindrical wave. The cylindrical wave is produced by focusing a plane wave with a conventional cylindrical lens. The distance between the film plane and the focal plane for the cylindrical wave (i.e., the focal length of the holographic lens being recorded) is selected by moving the cylindrical lens relative to the film. Both the film and the cylindrical lens are mounted on mini-computer controlled translation stages. A facet is exposed via a computer controller shutter. To record the next facet, the film is translated by one facet height and the cylindrical lens repositioned to achieve the prescribed focal length for this facet. The process is repeated until all facets are exposed. The final hologram consists of 13 facets each approximately 19×3mm in size.

The experimental results in Figs. 5 and 6 show the utility of our system. Figure 5 shows the focal pattern obtained when a conventional cylindrical lens is used to focus a line on a 5.4cm diameter sphere. The top of the sphere is located one focal length from the cylindrical lens. Figure 6 shows the improved results when our multi-facet optical element is used. The focal line is now sharp and the segmenting barely visible.

IMAGE FORMATION

Our last demonstration is in the use of multi-facet holograms for the formation of images. The hologram surface is again subdivided into numerous small facets, each of which contains a grating of a specified spatial frequency. When illuminated by a plane wave as in Fig. 7, the light is diffracted to specific locations in the image plane which is located a short distance (20cm) after the hologram plane. The image is thus built up of a number of small patches of light.

As before, each facet is interferometrically recorded under computer control. The computer translates the film behind a mask so that individual facets can be sequentially exposed, and deflects the recording beam to produce the requisite spatial frequencies.

This image production method produces results similar to those from kinoforms[3] in that all of the incident light is transformed into an image. The present production technique has the advantage, however, that one only needs to control a transverse spatial frequency instead of a longitudinal phase shift.

A resultant image from this type of hologram is shown in Fig. 8. Here a 181 facet hologram is used to form a simple image.

CONCLUSION

We have demonstrated three types of multi-facet holographic optical elements which combine optical flexibility, high efficiency, and large space-bandwidth products. These elements should prove useful for a large number of custom optical applications.

REFERENCES

1. S. K. Case, P. R. Haugen, and O. J. Lokberg, "Multifacet Holographic Optical Elements for Wavefront Transformation," Appl. Opt. 20, 2670 (1981).

2. S. K. Case and P. R. Haugen, "Partitioned Holographic Optical Elements," Opt. Engin., March 1982.

3. B. Lesem, P. M. Hirsh, and J. A. Jordan, Jr., "The Kinoform: A New Wavefront Reconstruction Device," IBM J. Res. Develop. 13, 150 (1969).

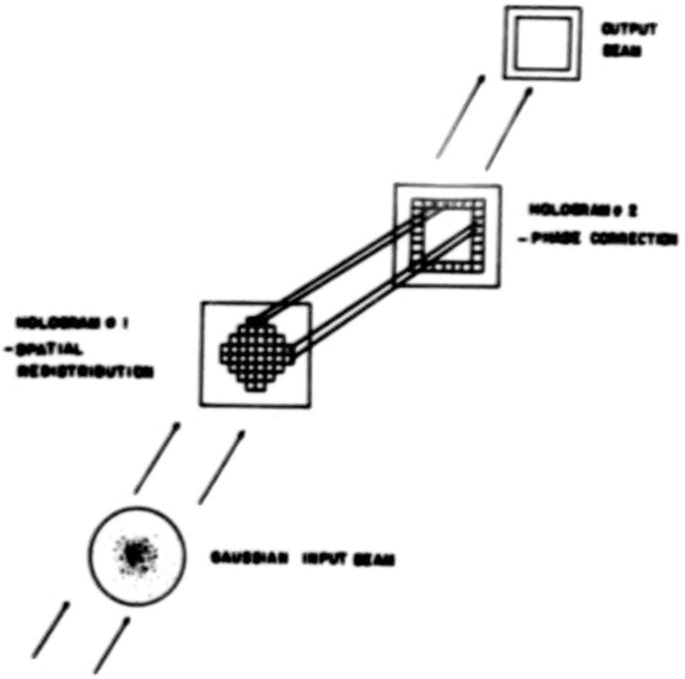

OUTPUT
BEAM
HOLOGRAM # 2
- PHASE CORRECTION
HOLOGRAM # 1
- SPATIAL
REDISTRIBUTION
GAUSSIAN INPUT BEAM

Figure 1.- Wavefront transformation via two holograms.

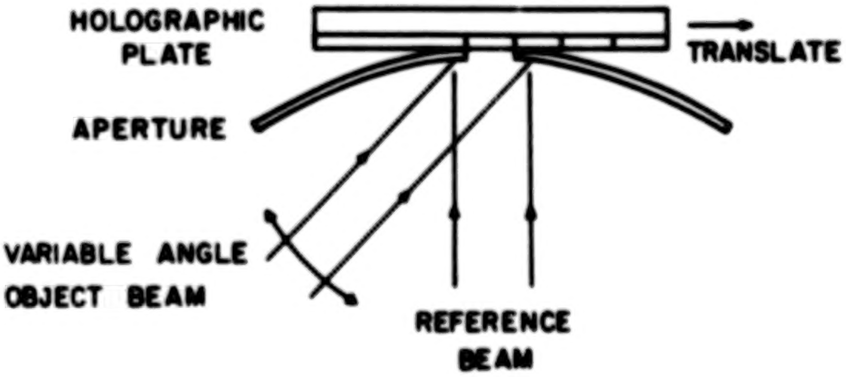

HOLOGRAPHIC
PLATE
TRANSLATE
APERTURE
VARIABLE ANGLE
OBJECT BEAM
REFERENCE
BEAM

Figure 2.- Multi-facet hologram recording method.

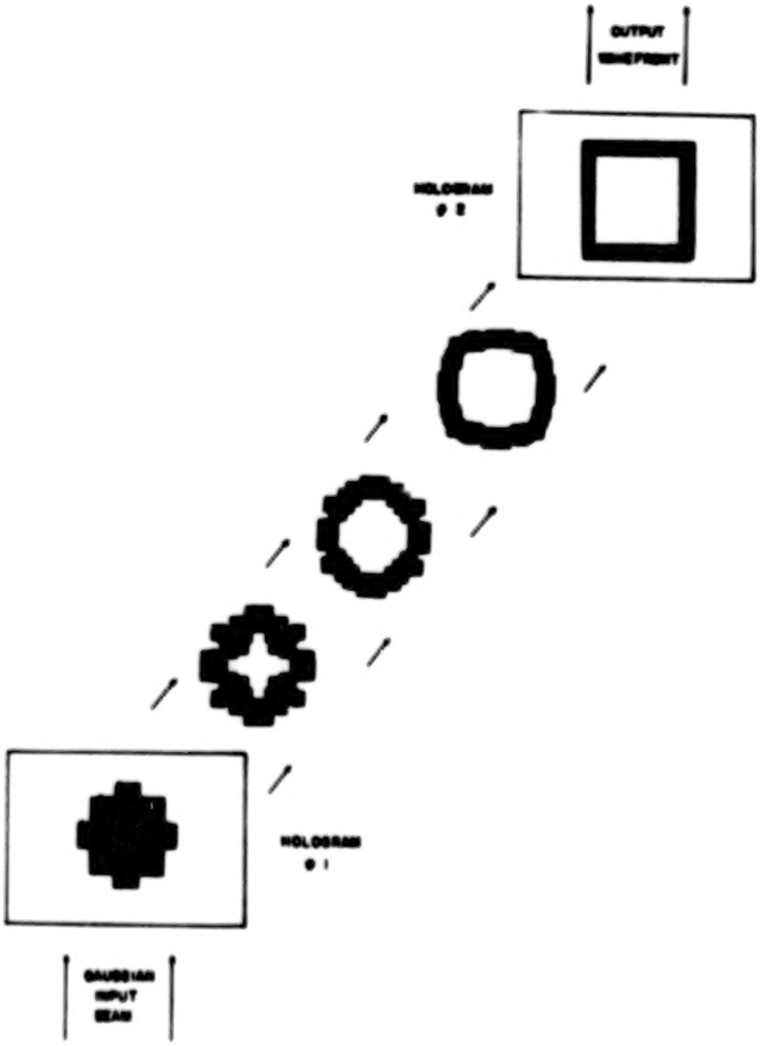

OUTPUT
WAVEFRONT
HOLOGRAM
# 2
HOLOGRAM
# 1
GAUSSIAN
INPUT
BEAM

Figure 3.- Light redistribution.

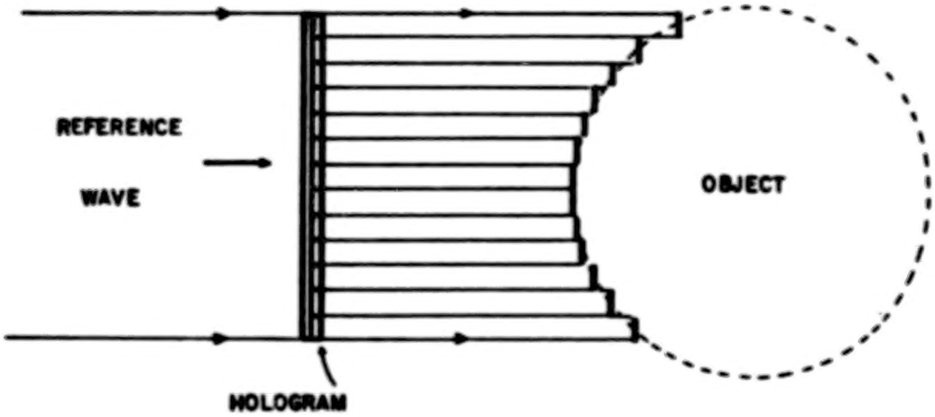

REFERENCE
WAVE
OBJECT
HOLOGRAM

Figure 4.- Side view of multi-facet hologram producing segmented focus on object surface.

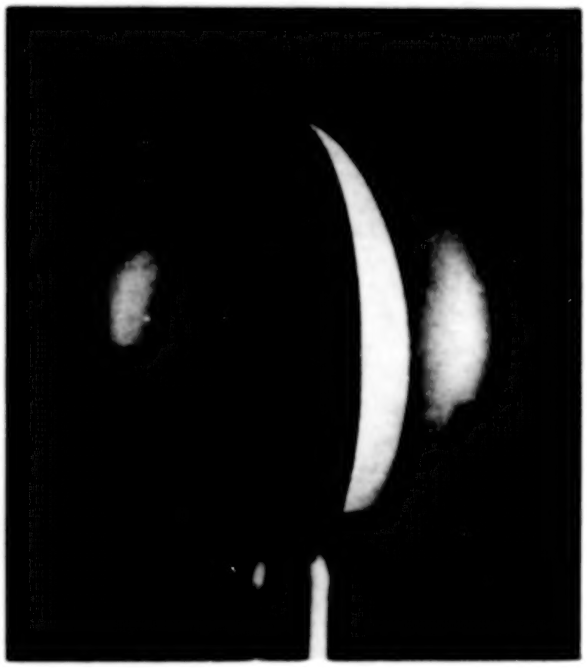


Figure 5.- Line focus formed by conventional cylindrical lens.

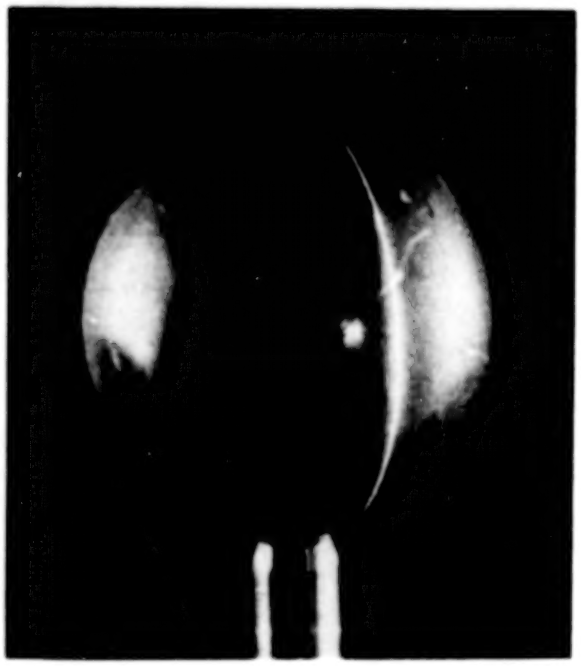


Figure 6.- Line focus formed by multi-facet holographic cylindrical lens.

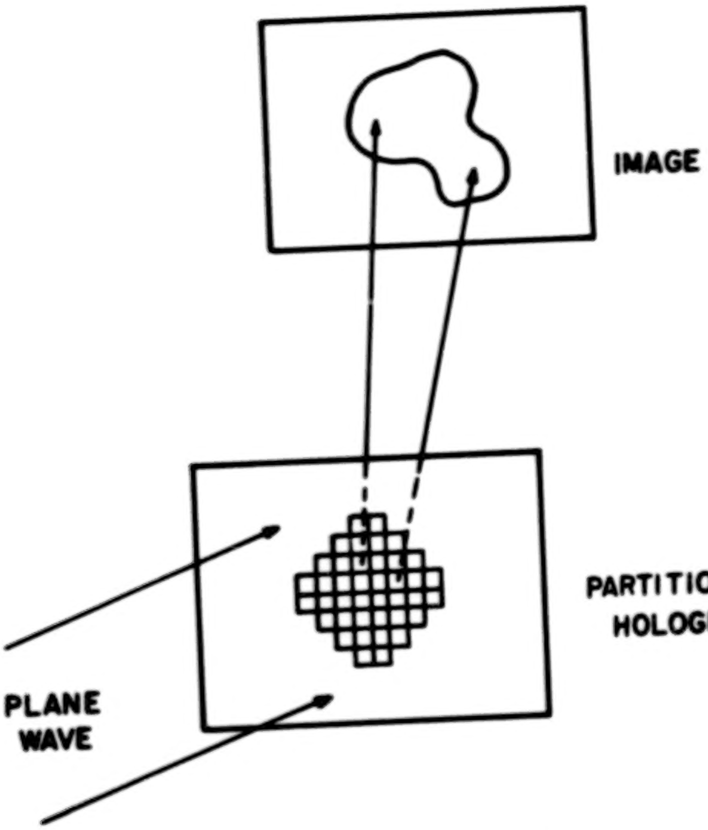


IMAGE
PARTITIONED HOLOGRAM
PLANE WAVE

Figure 7.- Image formation.

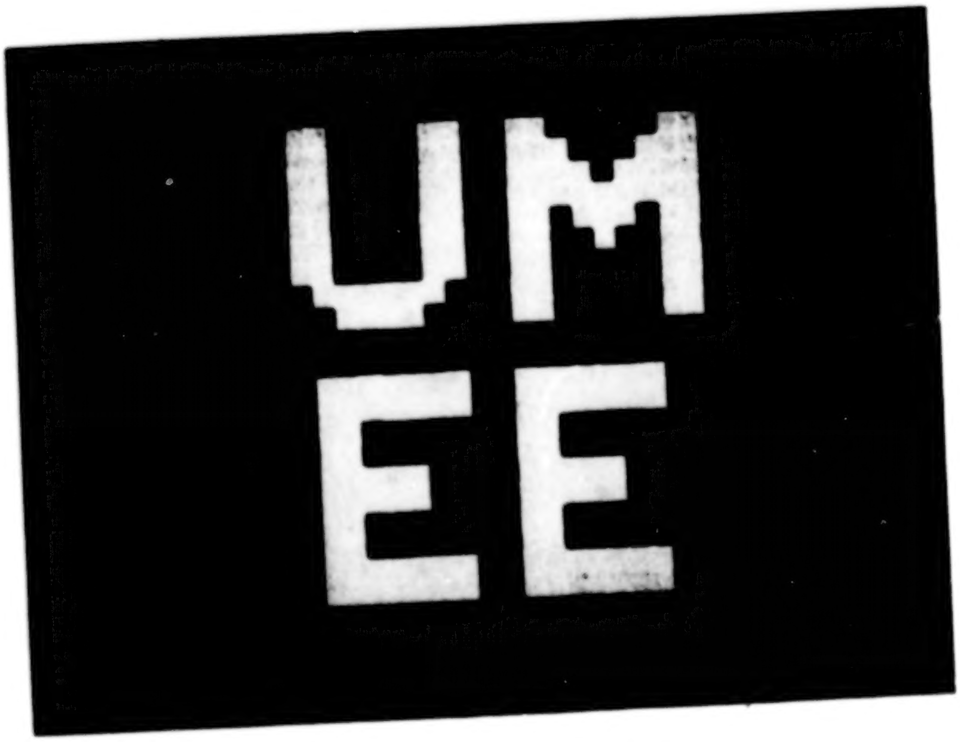


Figure 8.- Output image.

SEMICONDUCTOR LASER AND DETECTOR BANDWIDTH LIMITATIONS IN OPTICAL INFORMATION PROCESSING*

Sverre T. Eng
Chalmers University of Technology
S-412 96 Göteborg, Sweden

Semiconductor lasers and photodetectors are becoming important in compact time-integrating optical signal processing, optical systems performing arithmetic operations in parallel on arrays of binary data, fiber-optic data buses in hybrid optical/digital image processing systems, fast optical logic and detection for optical computers, and high-speed digital to analog converters, etc.

The limitations in bandwidth of the semiconductor laser transmitter as well as the photodetection receiver are of utmost importance in evaluating advanced real-time optical processing system potentials. Therefore, the transmission limitation has been investigated in a general way by PCM fiber-optic link studies at the highest possible (5 Gbit/s) bit rate.

A 5-Gbit/s optical fiber communication system has been built using TJS lasers and 3-Gbit/s transmission was obtained with a BH laser. A 500 m single-mode fiber was used. The laser modulation was limiting the bit rate. The rest of the system was capable of at least 10 Gbit/s. Bit error rates of 10^{-9} were measured and proved the feasibility of the system. The transmitter consisted of three parts: a multiplexer, a laser driver circuit, and a laser. In this experiment channels with a bit rate of 100 Mbit/s were multiplexed to a maximum bit rate of 5 Gbit/s. Three channels were used to carry information and one to transmit the clock signal to the receiver. The 100-Mbit/s input bit stream can have a nonreturn-to-zero of return-to-zero data format. The receiver consisted of an avalanche photodiode (BPW 28 from Telefunken), a MESFET low-noise preamplifier, and a demultiplexer. At a transmission path following the receiver, AND gates are mounted with a spacing equal to the inverted bit rate. The AND gate has a switching speed in the 10-Gbit/s range and consists of GaAs dual-gate MESFET transistors.

---

* Because this paper was not available at the time of publication, only the abstract is included.

Optical transmission at 8 Gbit/s has been performed earlier with TJS lasers (ref. 1). However, 5 Gbit/s may be a better practical choice at the present time.

REFERENCE

1. R. Tell and S. T. Eng, "8 Gbit/s Optical Transmission with TJS GaAlAs Laser and p-i-n Detection," Electronics Letters, Vol. 16, No. 13, June 1980.

A LIQUID CRYSTAL ADAPTIVE LENS

Stephen T. Kowel*
Syracuse University
Department of Electrical and Computer Engineering
Syracuse, N.Y.

Dennis S. Cleverly
General Electric Company
Aircraft Equipment Division
Utica, N.Y.

SUMMARY

Creation of an electronically controlled liquid crystal lens for use as a focusing mechanism in a multi-element lens system or as an adaptive optical element is analyzed. Varying the index of refraction is shown to be equivalent to the shaping of a solid refracting material. Basic characteristics of liquid crystals, essential for the creation of a lens, are reviewed. The required variation of index of refraction is provided by choosing appropriate electrode voltages. The configuration required for any incoming polarization is given and its theoretical performance in terms of modulation transfer function derived.

INTRODUCTION

Lenses play a critical role in almost all optical processing systems. From the crude and inexpensive plastic lens to the diffraction-limited lens, their role in the optical world is unique. The focusing of a lens is always accomplished by mechanical motion of one portion of the lens relative to the desired focal plane or correspondingly to other elements of the lens. This mechanical movement, to be precise, requires sophisticated and expensive mechanical hardware.

This paper addresses the electronically controlled creation of a lens. Since the actual focusing action is provided by electrode voltages, the image formation can be controlled. Also, since the index of refraction is controlled point-to-point on the lens, adaptive performance can be provided as well.

---

* Supported by the U.S. Army Night Vision and Electro-Optics Laboratory, Fort Belvoir, Virginia, Contract No. DAAK-70-80-C-0053.

SYMBOLS

f	focal length (meters)
k	wave number (meters^{-1})
L	diameter of lens (meters)
$n(x,y)$	position variable index of refraction
n_e	extraordinary index of refraction
n_o	ordinary index of refraction
n_θ	angle variable index of refraction
R_1	radius of curvature (meters)
$U(x,y)$	field distribution
V_T	threshold voltage (volts)
Δ	cell thickness (meters)
$\epsilon_\perp$	component of the dielectric tensor perpendicular to the molecular axis
$\epsilon_\parallel$	component of the dielectric tensor parallel to the molecular axis

LENS CREATION BY INDEX OF REFRACTION VARIATION

Lenses are classically created by varying the optical path length over an aperture by radially shaping a refractive medium. The same phase transformation can be obtained if the material has parallel faces but a varying index of refraction.

Consider the geometry shown in figure 1, where a material of width Δ and variable index of refraction (ref. 1)

$$n(x,y) = n_e - \frac{r^2}{2\Delta f} \qquad r^2 = x^2 + y^2 \tag{1}$$

is centered on the z-axis and lies in the x-y plane. Making the same approximation as in a phase analysis of a thin lens (ref. 2), namely,

1) Light passing through the device suffers only a phase transformation
2) Light rays are paraxial, that is,

$$\sqrt{1 - \frac{x^2 + y^2}{R_1^2}} \simeq 1 - \frac{x^2 + y^2}{2R_1^2} \tag{2}$$

results in a transmission function

$$t(x,y) = \exp\left(\frac{-jk}{2f}\,(x^2 + y^2)\right) \tag{3}$$

where

$$f \triangleq \frac{n_e - 1}{R_1} \tag{4}$$

LIQUID CRYSTAL CHARACTERISTICS

All crystals under the influence of an electric field or applied stress exhibit birefringence (ref. 3). That is, the index of refraction can be varied. In solids, however, this effect is small since only distortions of the indicatrix are made. Liquid crystals are unique since the molecules reorient under the influence of an applied field, either electric or magnetic (ref. 4). In addition, nematic liquid crystals are uniaxial and the optical axis is coincident with the molecular axis. Thus, as the molecules reorient in response to an applied field, the entire indicatrix is rotated, making large changes in the index of refraction.

Liquid crystals can be divided into two groups. If the dielectric tensor is such that the component along the molecular axis ($\varepsilon_{\parallel}$) is greater than the component perpendicular to the axis ($\varepsilon_{\perp}$), the crystal is said to be positive. These molecules tend to align parallel to an applied field. If the reverse is true, $\varepsilon_{\parallel} < \varepsilon_{\perp}$, the crystal is said to be negative. In this case, the molecules will align perpendicularly to an applied field. Both types of liquid crystals are positive uniaxial crystals. The extraordinary index of refraction, n_e (along the optical axis which is coincident with the molecular axis), is greater than the ordinary index of refraction, n_o.

Figure 2 illustrates the changing of index of refraction with orientation of the liquid crystal molecules. In this example, a liquid crystal is shown at the top of the figure in the homogeneous state. Correspondingly, the indicatrix is aligned so that a wave polarized in the x-direction is influenced by an index of refraction n_e. At the bottom of the figure, the orientation has been altered to the homeotropic state, in which case the same polarized light encounters an index of refraction n_o. At an intermediate point, the index of refraction is dependent on θ, the direction of the nematic director, and is given by (ref. 1)

$$\frac{1}{n_\theta^{\,2}} = \frac{\cos^2\theta}{n_e^{\,2}} + \frac{\sin^2\theta}{n_o^{\,2}} \tag{5}$$

Two important points should be noted. First, only the light polarized in the x-direction was influenced by a changing index of refraction. Light polarized in the y-direction suffered a constant phase delay. Because independent of the molecular orientation in the x-z plane, it was influenced by a constant index of refraction, n_o. The second important point is that the liquid crystal molecules do not align uniformly across a cell. As the voltage is increased above threshold (ref. 5), the molecules near the center react to the field first. At saturation, all of the molecules have responded except those tied by boundary conditions at the cell wall.

Figure 3 illustrates the index of refraction variation with voltage in one particular material.

LIQUID CRYSTAL LENS

Figure 4 shows a single stage of a liquid crystal lens. With voltages appropriately set on the electrode array, light polarized in the x-direction would be given the correct phase transformation. The y-polarized component would suffer a constant phase delay. To obtain total polarization capability, there must be a second stage with the nematic director having a preferential turn in the y-direction. This stage would vary the index of refraction for the y-component and give a constant phase delay to the component polarized in the x-direction.

It has been assumed to this point that electrodes could be configured to give the required radial variation in index of refraction. Since the voltage variation desired is radial, it would appear that radially symmetric electrodes would be ideal. However, radially symmetric electrodes require difficult mask fabrication and connections made to them create obscurations. Row and column addressable, rectangular grid electrodes like those used in commercially available liquid crystal displays could be considered. In this case, addressing individual electrodes is easy and mask fabrication is a simple process, but the device inherently does not match the symmetry of a lens. Thus, the structure would cause aberrations.

A simple structure can be implemented by using crossed linear arrays in tandem. Consider a square aperture of width L centered at the origin. Let the transmission of the aperture be given by

$$t(x,y) = \text{rect}\left(\frac{x}{L}\right)\text{rect}\left(\frac{y}{L}\right)\exp\left(jk\Delta n(x)\right) \tag{6}$$

where

k = wave number

Δ = thickness of the material

and

$n(x) = n_e - \frac{x^2}{2\Delta f}$ = index of refraction

Just past the aperture, if it is illuminated by a unit-amplitude, normally incident plane wave, the field distribution U' is

$$U'(x,y) = \text{rect}\left(\frac{y}{L}\right)\text{rect}\left(\frac{x}{L}\right)\exp\left(jk\Delta\left(n_e - \frac{x^2}{2\Delta f}\right)\right) \tag{7}$$

Ignoring the constant phase term,

$$U'(x,y) = \text{rect}\left(\frac{y}{L}\right)\text{rect}\left(\frac{x}{L}\right)\exp\left(\frac{-jkx^2}{2\Delta f}\right) \tag{8}$$

Now consider placing another transmission factor directly in front of the aperture with

$$n(y) = n_e - \frac{y^2}{2\Delta f} \tag{9}$$

then

$$U' = t_1(x,y)\, t_2(x,y) \tag{10}$$

$$U' = \text{rect}\left(\frac{y}{L}\right) \text{rect}\left(\frac{x}{L}\right) \exp\left(\frac{-jk(x^2 + y^2)}{2f}\right) \tag{11}$$

the ideal form of a lens is obtained with linear electrodes. By cascading two stages not only is the mask fabrication process simplified, the required symmetry is also maintained.

Figure 5 shows a single stage cross section and a representative electrode structure. Note that the ground electrode is uniform. Also, the variation in voltage profile can be used to generate either a change in index of refraction for x or y polarized light. If the preferential direction of the liquid crystal director is in the x-direction, only the component of the light polarized in that direction will be affected by a varying index of refraction.

To obtain a complete lens, four stages are required. "Complete" means that any incoming polarization is given the appropriate phase transform of a thin lens. Figure 6 illustrates the functions of the four stages. The first stage gives the appropriate variation in index of refraction in the x-direction for light polarized in the x-direction. Similarly, the second stage varies the index of refraction in the y-direction but still only for light polarized in the x-direction. Thus, if a lens was to be created for x-polarized light only, the first two stages would be sufficient. The last two stages repeat the process for the component polarized in the y-direction. Figure 7 shows the cross section of the four-stage liquid crystal lens.

OPTICAL PERFORMANCE ANALYSIS

Due to the electrode structure used to create the index of refraction variation in the liquid crystal, a diffraction-limited lens will not be formed. The wavefront leaving the lens will have phase distortions caused by the approximation of a smoothly varying index of refraction by a sampled function. Thus, the liquid crystal lens will inherently have aberrations. However, the smaller the electrodes and the spacing between them, the better the index of refraction is matched to a smooth curve and therefore the more nearly the lens approaches the diffraction limit. Figures 8 and 9 show the predicted MTF for a 2-cm wide lens with a 10-meter focal length constructed with 501 and 101 electrodes. Spaces between electrodes are assumed to be the same size as the electrodes. Thus, fabrication of the 501 electrode (20-micron width) is easily achievable with today's microelectronics capability.

CONCLUSION

An electronically controlled liquid crystal lens will provide direct control on focusing and point-to-point control on index of refraction for potential adaptive optical techniques. The lens has all the favorable characteristics of liquid crystal displays, such as operation at low voltages and low power dissipation. With microprocessor control, it could adapt in real time and be electronically calibrated. Using current microelectronic technology, near diffraction-limited performance is predicted.

REFERENCES

1. Kowel, S.T., et al.: Polymeric Microelectronics. Annual Technical Report prepared for U.S. Army Night Vision and Electro-Optics Laboratory, Fort Belvoir, Virginia. Contract No. DAAK-70-80-C-0053.

2. Goodman, Joseph W.: Introduction to Fourier Optics. McGraw-Hill, New York, 1968.

3. Nye, J.F.: Physical Properties of Crystals. Oxford Press, London. 1972.

4. Priestley, E.B., et al.: Introduction to Liquid Crystals. Plenum Press, New York. 1974.

5. Gruler, H., and Meier, G.: "Electric Field Induced Deformation in Oriented Liquid Crystals of the Nematic Type," Mol. Cryst. and Liq. Cryst. vol. 16, 1972, p. 299.

BEST DOCUMENT AVAILABLE

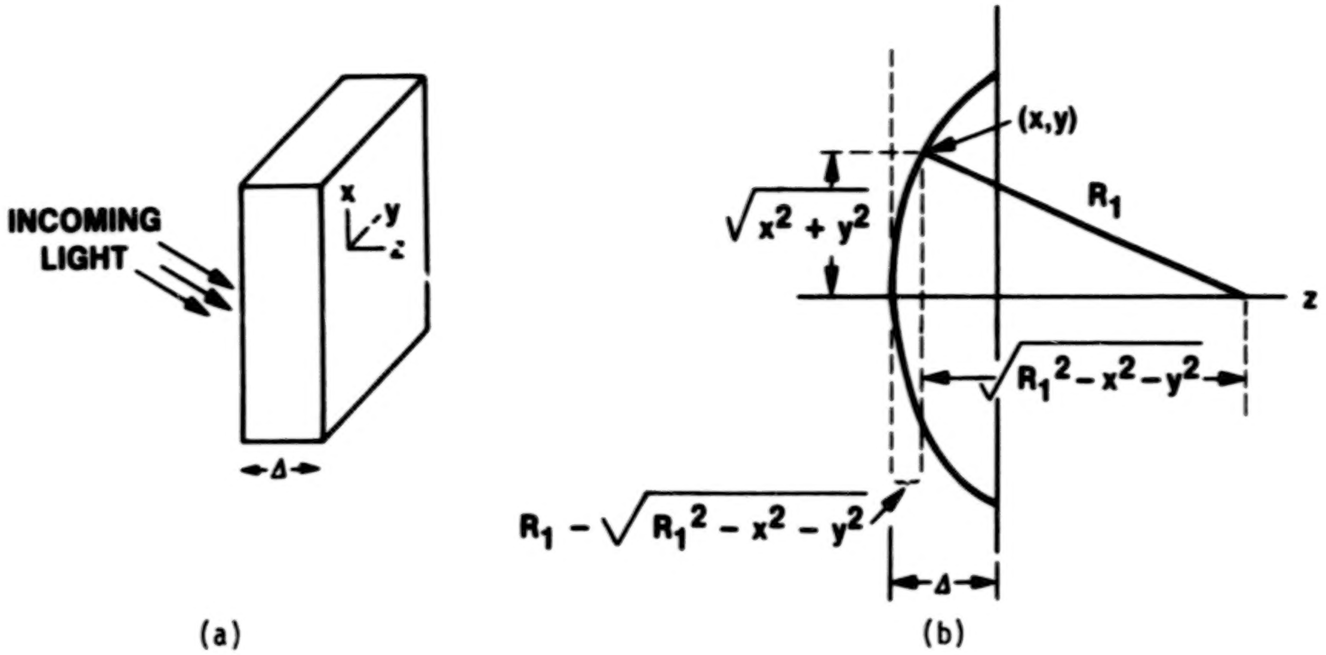

INCOMING LIGHT
x
y
z
Δ
(x,y)
R1
$\sqrt{x^2+y^2}$
z
$\sqrt{R_1^2-x^2-y^2}$
$R_1-\sqrt{R_1^2-x^2-y^2}$
Δ
(a)
(b)

Figure 1.- Basic configuration (a) and geometry of variation (b).

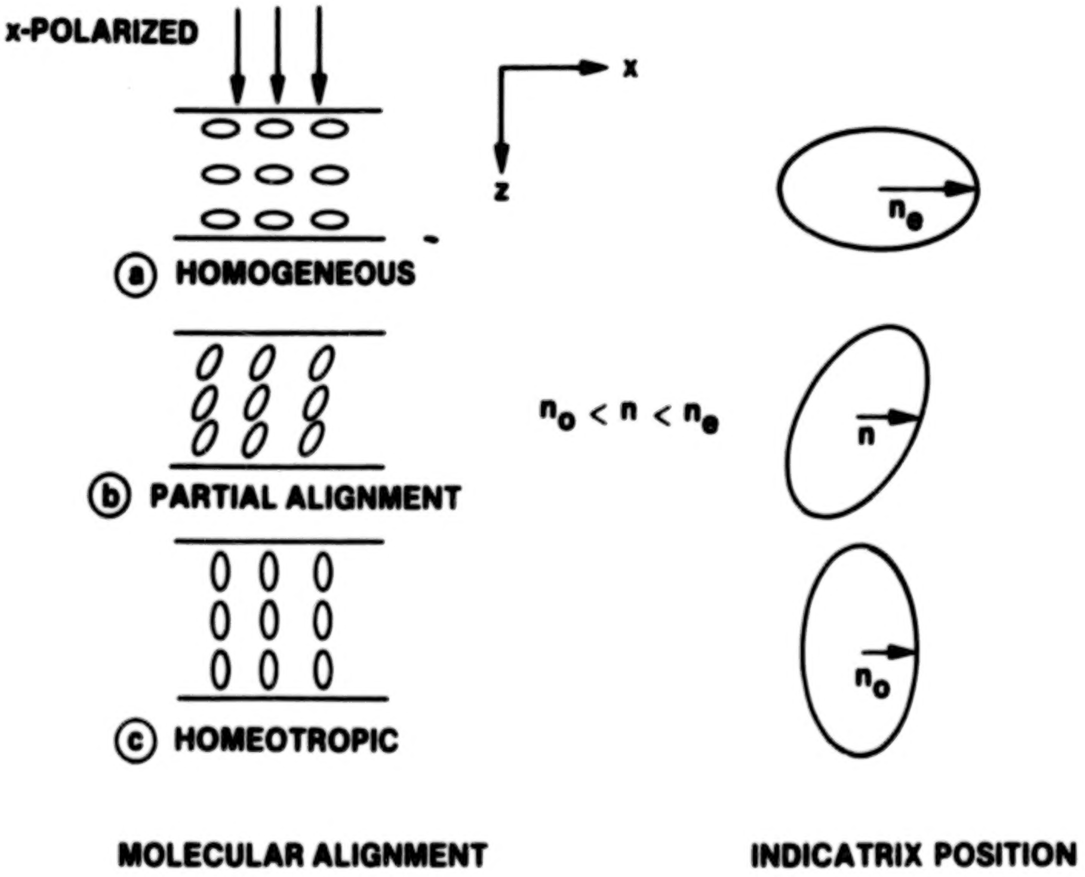

x-POLARIZED
x
z
ⓐ HOMOGENEOUS
ⓑ PARTIAL ALIGNMENT
ⓒ HOMEOTROPIC
n_e
$n_o < n < n_e$
n
n_o
MOLECULAR ALIGNMENT
INDICATRIX POSITION

Figure 2.- Liquid crystal orientation for variation of the index of refraction.

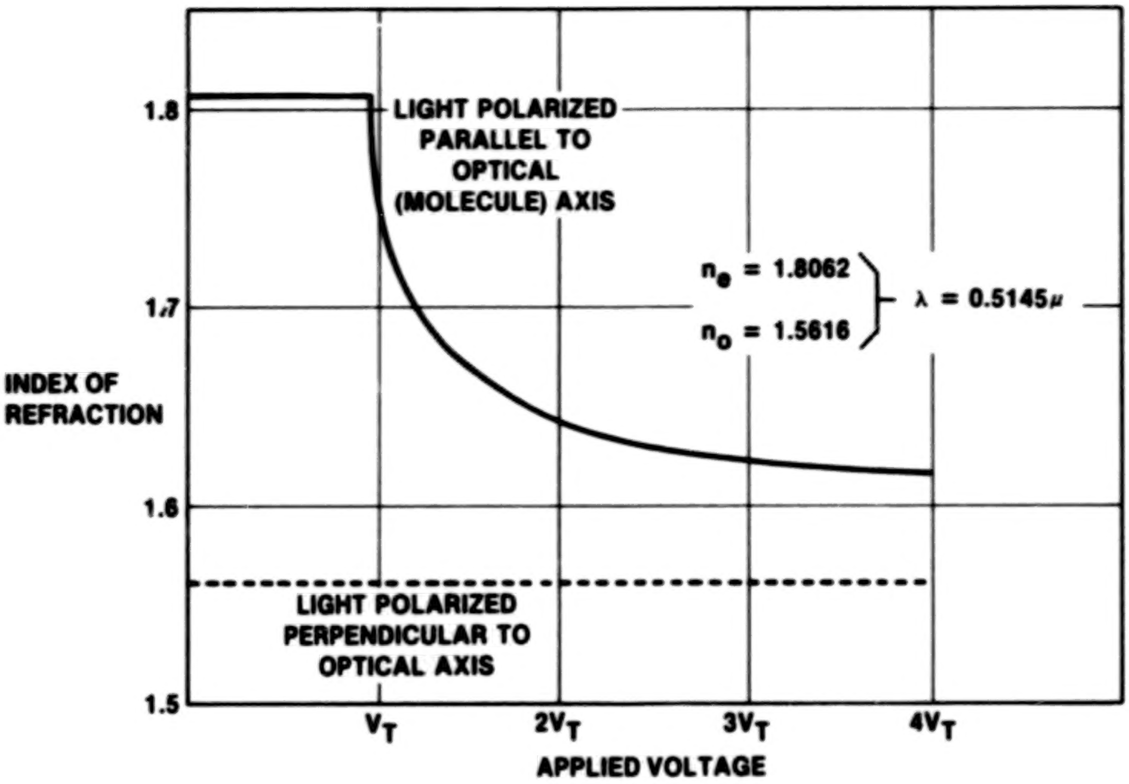


Figure 3.- Index of refraction versus cell voltage using MBBA.

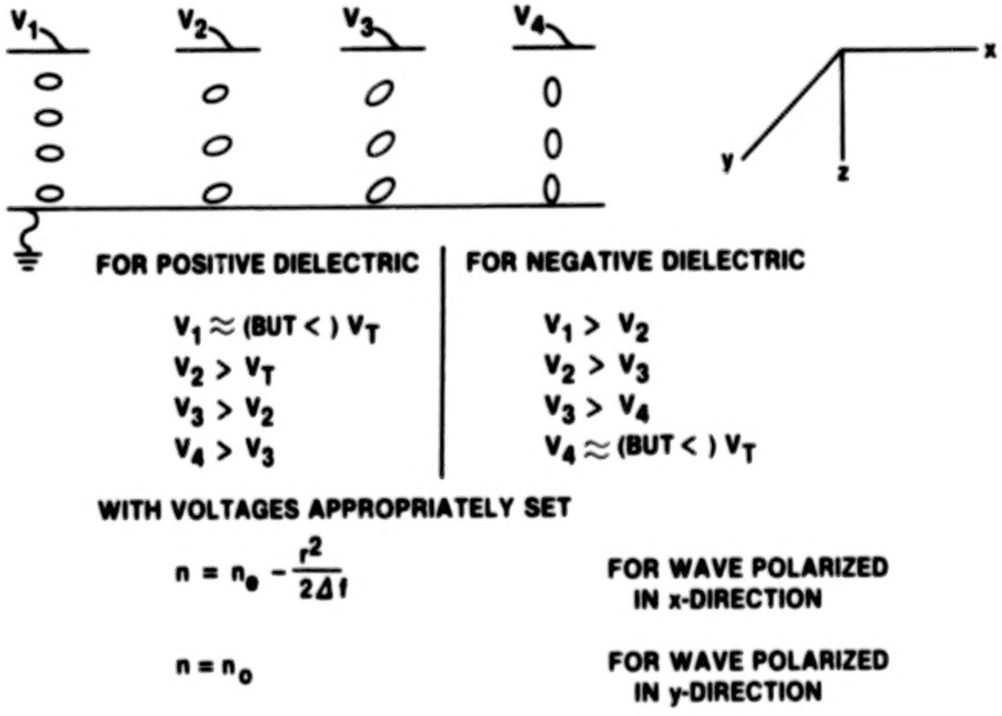

V_1
V_2
V_3
V_4
x
y
z
FOR POSITIVE DIELECTRIC
$V_1 \approx$ (BUT $<$) V_T
$V_2 > V_T$
$V_3 > V_2$
$V_4 > V_3$
FOR NEGATIVE DIELECTRIC
$V_1 > V_2$
$V_2 > V_3$
$V_3 > V_4$
$V_4 \approx$ (BUT $<$) V_T
WITH VOLTAGES APPROPRIATELY SET
$n = n_e - \frac{r^2}{2\Delta f}$
FOR WAVE POLARIZED IN x-DIRECTION
$n = n_o$
FOR WAVE POLARIZED IN y-DIRECTION

Figure 4.- Cross section of liquid crystal lens.

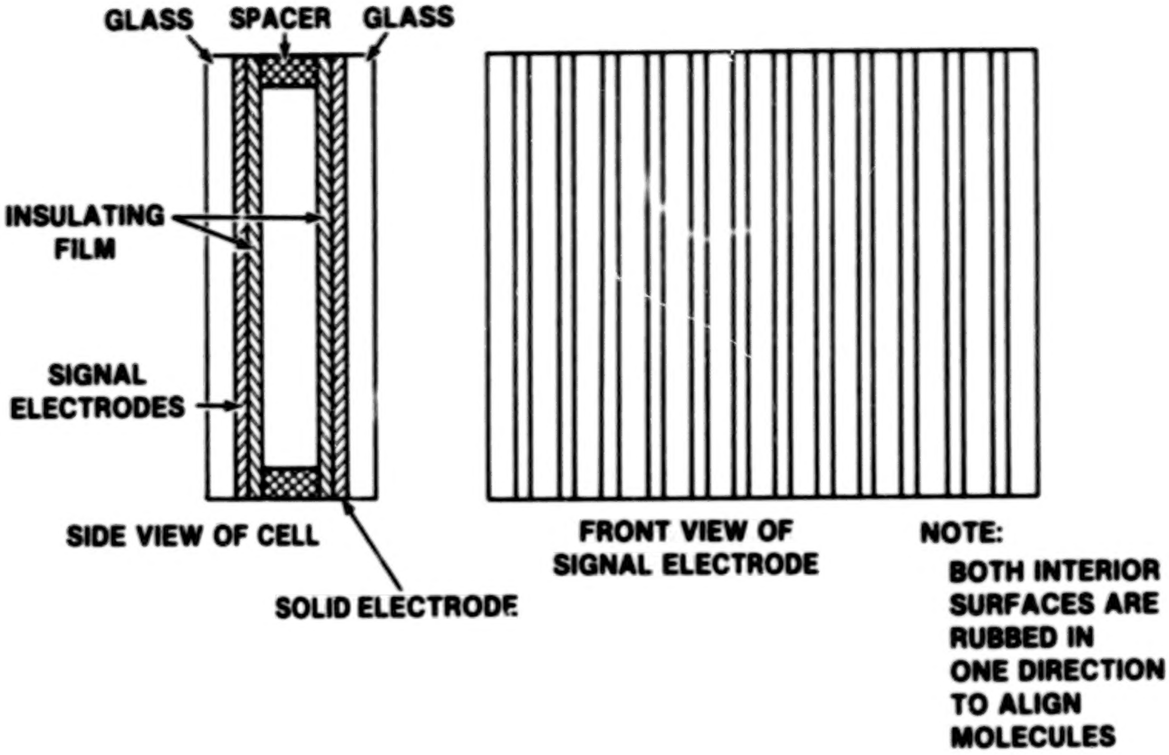

GLASS
SPACER
GLASS
INSULATING FILM
SIGNAL ELECTRODES
SIDE VIEW OF CELL
SOLID ELECTRODE
FRONT VIEW OF SIGNAL ELECTRODE
NOTE: BOTH INTERIOR SURFACES ARE RUBBED IN ONE DIRECTION TO ALIGN MOLECULES

Figure 5.- Single stage of liquid crystal lens.

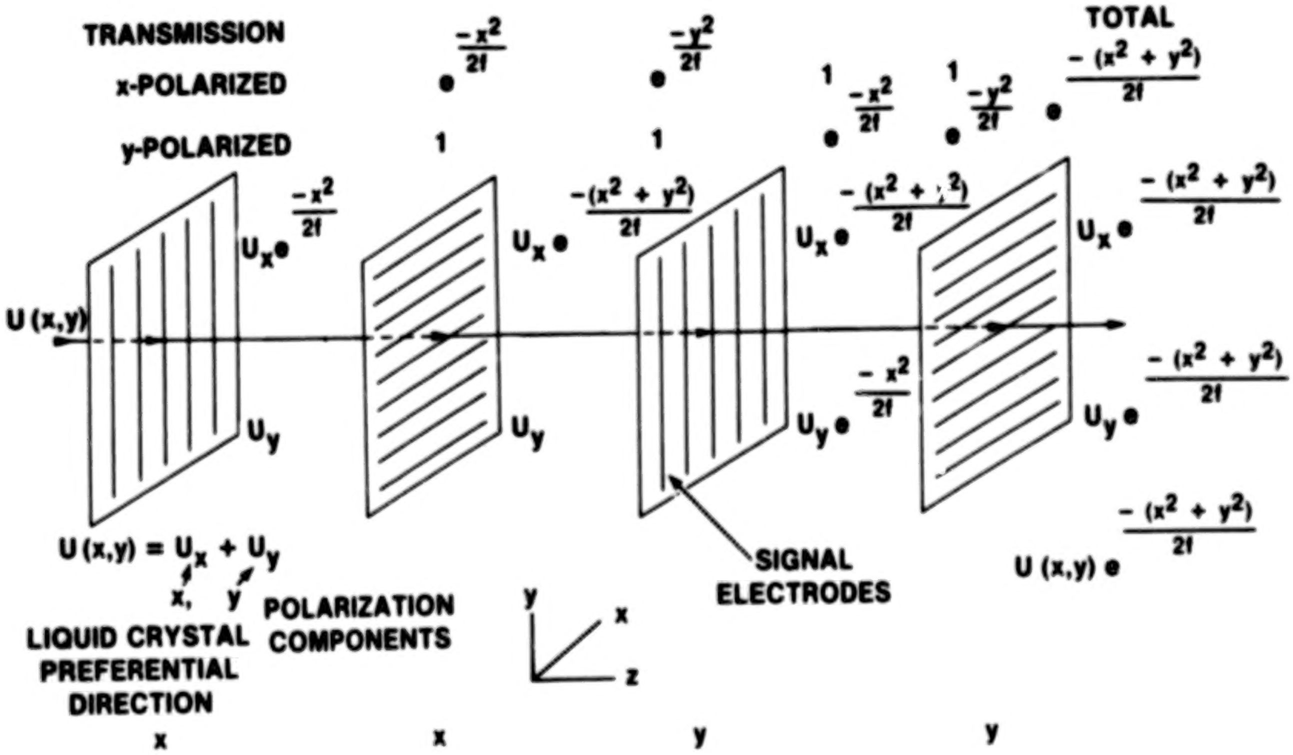

TRANSMISSION
x-POLARIZED
y-POLARIZED
TOTAL
$e^{\frac{-x^2}{2f}}$
$e^{\frac{-y^2}{2f}}$
1
1
1
1
$e^{\frac{-x^2}{2f}}$
$e^{\frac{-y^2}{2f}}$
$e^{\frac{-(x^2+y^2)}{2f}}$
$U(x,y)$
$U_x e^{\frac{-x^2}{2f}}$
U_y
$U_x e^{\frac{-(x^2+y^2)}{2f}}$
U_y
$U_x e^{\frac{-(x^2+y^2)}{2f}}$
$U_y e^{\frac{-x^2}{2f}}$
$U_x e^{\frac{-(x^2+y^2)}{2f}}$
$U_y e^{\frac{-(x^2+y^2)}{2f}}$
$U(x,y)\, e^{\frac{-(x^2+y^2)}{2f}}$
$U(x,y) = U_x + U_y$
x, y
POLARIZATION COMPONENTS
LIQUID CRYSTAL PREFERENTIAL DIRECTION
SIGNAL ELECTRODES
y
x
z
x
x
y
y

Figure 6.- Functions of the four stages in a liquid crystal lens.

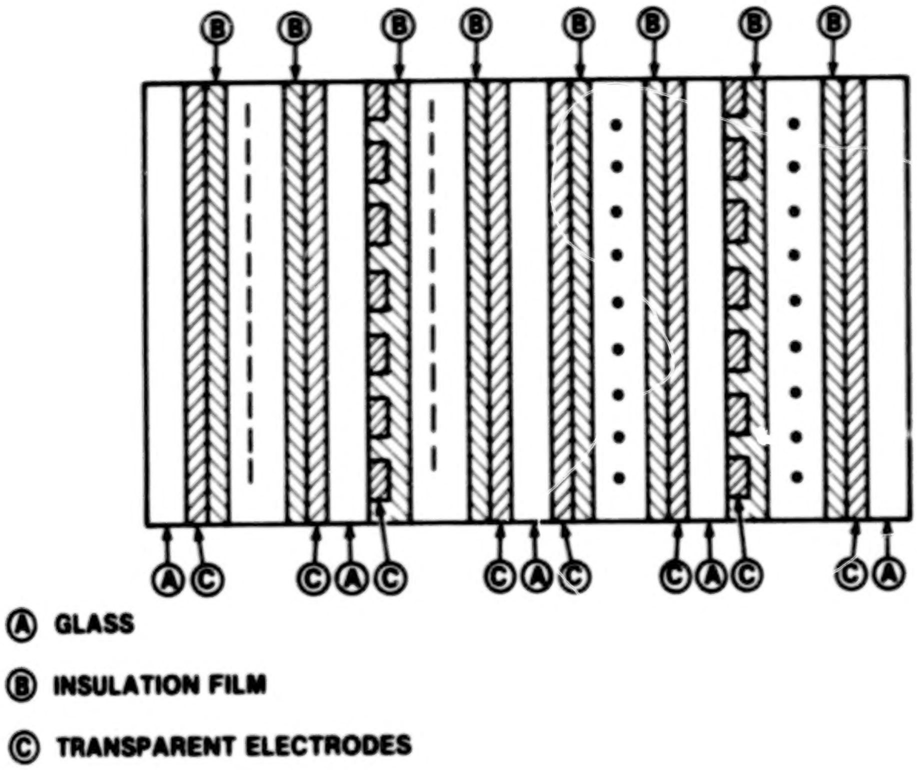

Ⓑ
Ⓐ
Ⓒ
Ⓐ GLASS
Ⓑ INSULATION FILM
Ⓒ TRANSPARENT ELECTRODES

Figure 7.- Four-stage liquid crystal lens.

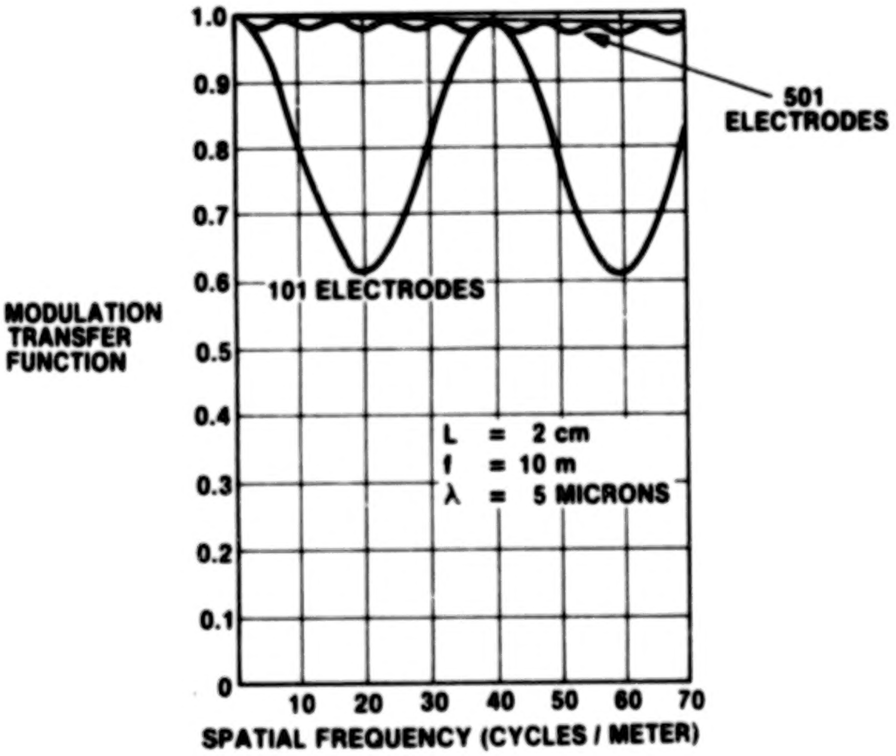


Figure 8.- Modulation transfer function of liquid crystal lens. λ=5.0 μm.

BEST DOCUMENT AVAILABLE

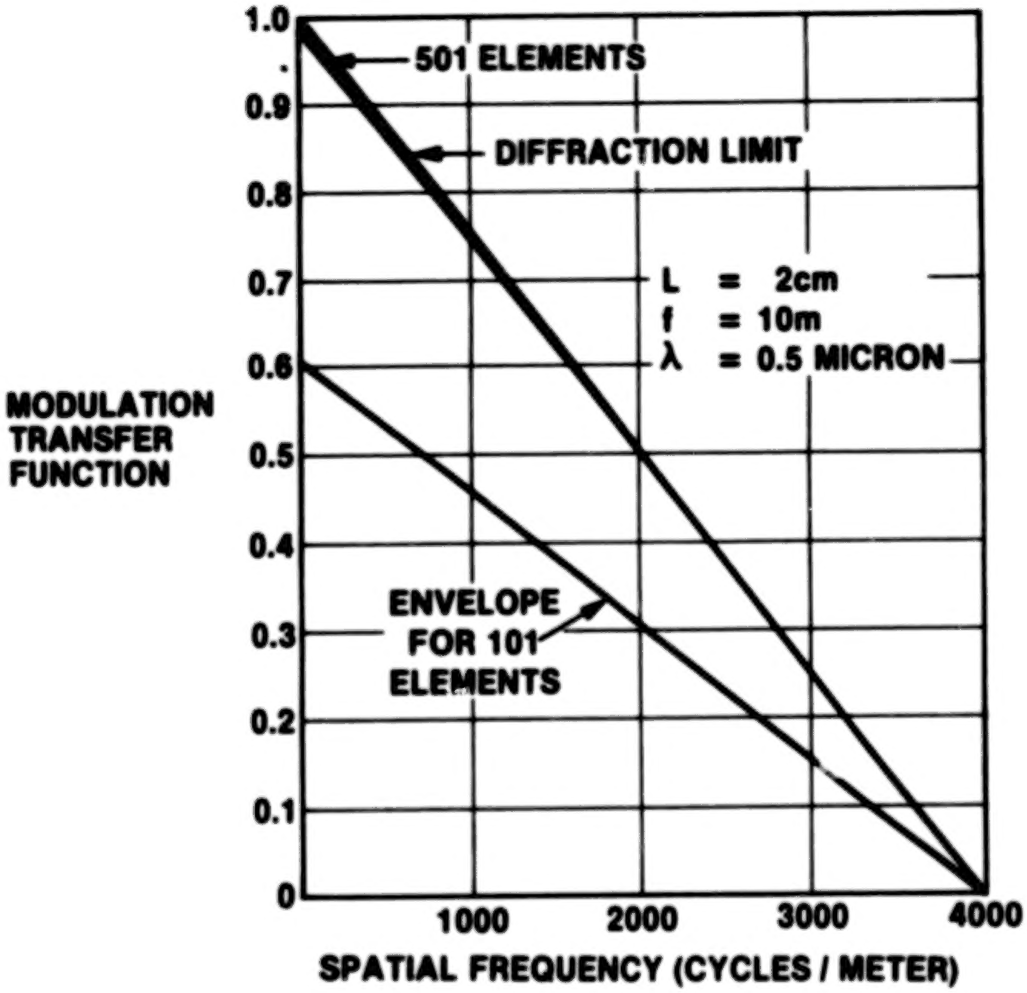


Figure 9.- Modulation transfer function of liquid crystal lens. λ=0.5 μm.

MANIPULATION OF LIGHT WITH MAGNETO-OPTIC STRIPE DOMAIN FILMS

E. J. Torok, F. G. Hewitt, J. A. Krawczak, G. L. Nelson

Sperry Univac, DSD, St. Paul, Minnesota

Work sponsored in part by the Army Research Office (ARO), Research Triangle Park, North Carolina.

ABSTRACT

Magnetic diffraction grating materials are currently being developed to provide a simple means of deflecting light in a two-dimensional, solid-state fashion. The most promising material, for several applications, appears to be bismuth substituted iron garnet films in epitaxial form. Calculations indicate that deflection efficiency greater than 60% is possible in the near-infrared region of the spectrum. Within the field of view of the deflector, measurements predict that 10^5 resolvable spots can be expected. Applications include 1) general purpose deflection of free laser light, 2) image processing of extended sources such as transparencies, 3) programmable lensing, and 4) fiber optic matrix switching.

MAGNETO-OPTIC LIGHT DEFLECTOR

The active component of this deflector is a dynamically alterable, solid-state phase diffraction grating that is the energetically favored domain structure of properly configured magnetic materials. A description of the deflector follows.

Diffraction of a light beam occurs as a result of periodic variations in the wave amplitude or phase across a wave normal surface. Magnetic stripe domain arrays can introduce a periodic 180° phase variation in an incident optical field, through magnetic birefringence. A stripe domain is a long, straight region of uniform width in which the magnetization is nearly constant. Typically, the width can vary from .5 micron to 30 microns or more while the length can extend to several centimeters.[1] Stripe domains in a given sample may or may not have equal width, depending on sample properties and on imposed magnetic fields.

Consider a linear array of stripe domains in a magnetic platelet as depicted in figure 1. Adjacent stripes have $\hat{z}$ components of magnetization that are antiparallel and usually have a continuous component in the x y plane. Because of the Faraday effect the E_x and E_y components of an incident optical field suffer clockwise rotation in odd-numbered stripes and counterclockwise rotation in even-numbered stripes. This differential rotation provides an electric component that is orthogonal to the incident polarization and has 180° alternations (parallel and antiparallel to $\hat{y}$) that match the spatial period of the domains. The process operates uniformly for all incident polarizations, including random.

In the far-field emerging light adds constructively at angles θ_n, given by

$$\sin \theta_n = \frac{n\lambda}{\Lambda} ,$$

where n is the order number, λ the incident wavelength, and Λ the grating period. In the special case where each stripe has the same width, even orders are suppressed. Magnetic apodization at the transitions between even and odd stripes discourages higher orders, as well.

The n^{th} odd-order power diffraction efficiency for a square phase grating is found from

$$\frac{I_n}{I_o} = \frac{4}{n^2\pi^2} e^{-\alpha(\lambda)t} \sin^2 F(\lambda)t ,$$

with α the optical absorption coefficient, t the material thickness, and F the Faraday rotation.[2] Figure 2 shows the potential total efficiency of several candidate crystalline deflector materials in the visible and infrared region of the spectrum, based on reported values for F and α.[3-6]

In order to alter the grating in a solid-state dynamic fashion, reliance is made on the strong coupling of the material magnetization and applied magnetic fields either coplanar with, or perpendicular to, the stripe domains. There are several ways in which a magnetic grating can be field programmed to deflect light. A perpendicular field changes the stripe width, periodicity, or both, causing linear deflection and perhaps a shuffling of light amongst the various allowed orders. If the field is applied in the plane, and collinear with the domains, then the stripe width varies with H^{-1}, increasing the diffraction angle with increases in field intensity. Finally, if the field is applied to a general direction in the plane, the grating is reestablished collinear to the field, resulting in azimuthal deflection. For some materials the field strength necessary to cause deflection in the annular field of view is less than that required for Lorentz deflection of electron beams in CRTs. Thus, microsecond switching speeds are possible with watt level electrical power.

GARNET

Of the deflector candidates contrasted in figure 2, rare earth iron garnet is presently the preferred material for a variety of reasons. It is readily obtainable as epitaxial films up to 100 microns thick grown on gadolinium gallium garnet, commonly employed as magnetic bubble memory substrates. Stable stripe domain arrays of sufficient quality to provide 10^5 resolvable spots in the annular field of view have been observed. Curie points to 550°K insure that the magnetization is nearly constant over a wide temperature range, including room ambient. The applied fields required to manipulate the stripe domains are on the order of 100 oe., derivable from computer controlled Helmholtz pairs or strip lines in close proximity to the crystal. With bismuth doping the Faraday rotation can reach 50,000°/cm in the visible and 12,000°/cm in the near-infrared.[7] The optical absorption exhibits a local minimum at the .81 micron fiber-optic wavelength and a large window of near-zero absorption at wavelengths greater than 1.2 microns. Calcium doping has been successfully used to reduce the absorption at important wavelengths.[8] From experimental data obtained with a Faraday hysteresigraph and a spectrophotometer on thin epitaxial samples it is possible to calculate the maximum deflection efficiency of thick samples or of thin samples operated inside an optical resonant cavity. The calculation has been done and the results are shown in figure 3. Work is currently under way to experimentally evaluate these expectations and to develop the necessary crystal growth facilities to exploit the potential of bismuth garnet for light deflector applications.

APPLICATIONS

Agile light deflectors based on magneto-optics offer unique solutions to a number of optical switching and processing problems. Obvious applications include laser radar at 1.06 microns and 10.6 microns, focal plane array scanners, and optical communications between moving platforms. Because stripe domain arrays can be plastically deformed by spatially varying fields, adaptive optic processing of images of extended objects can also be performed. In guided wave communications, bismuth garnet, in its present form, serves quite well as a high fanout fiber optic switch.

Adaptive Optics

In a stripe domain grating the periodicity, stripe width, and orientation need not be constants over the aperture. Spatially varying magnetic fields can be utilized to locally modify the deflection of incident light from pointlike sources or from extended sources such as collimated beams or transparencies. By properly tailoring the grating, dynamically alterable compressors or expanders and image rotators can be generated.

One-dimensional compressors or expanders are constructed from the domain pattern that has the basic spatial modulation seen in figure 4a. Both the stripe density and grating orientation are functions of local field. This would be useful for correcting image distortions.

Gratings obtained by converting the linear array of a film with some in plane magnetic component, into the radial array of figure 4b, provide a means for continuous distortionless image rotation if the grating has the appropriate radial stripe density gradient. A geometric optic analysis reveals that the gradient is such that the stripe density is inversely proportional to radius. Then the impressed image rotation angle, B, is found to be

$$B = \tan^{-1} \frac{s\lambda}{2c} ,$$

with s the distance to the post grating image plane and c a quantity that depends on field magnitude and stripe widths. In addition, the image experiences a magnification of sec B. Continuous rotation occurs with intensity changes in the field.

In order to initialize the grating to a radial mode it is necessary to bring a point pole into approximate contact with the film surface. The symmetry in the pole's field causes the grating to assume the desired form but with constant stripe density.

In practice the pole can be supplied by a polished ferrite needle. The inplane field that develops the proper stripe density was found to match the tangential field produced by an extended magnetic polepiece. Thus, it has been possible to observe both the rotation and the scaling properties of the radial domain image processor. Rotations to near 90° were recorded. Since the required polepiece is available in ferrite form, high speed rotation is expected.

Arrays formed from domains that are concentric annuli about a fixed center may act like a Fresnel zone plate; i.e., collimated light focuses on the zone axis when the domain widths satisfy the pertinent Fresnel relations. If the incident light impinges on just a sector of this array as in figure 4c, focusing occurs off the optic axis. In either case, dynamic control of the domains implies dynamic lensing. Most likely, garnet materials for this application are of the bubble type because they have only a perpendicular magnetization component.

Fiber Optic Switch

Useful deflection efficiency at the present fiber optic wavelengths, along with magnetic control of the intrinsic grating in garnet films is the basis for a multiport fiber optic switchboard as seen in figure 5. Information-bearing light, propagating in any or all of the elements of the input fiber 2-dimensional matrix is collimated by gradient index lenses (GRIN). Light from a given fiber lens falls on just one deflector element which steers it to one fiber in the focal plane of the output lens. The switchboard has high fanout with just a single level of optical switching. Greater than 100 X 100 input or output arrays can be accommodated.

Figure 6 shows a basic version of a fiber optic switch that utilizes only the inner circumference of the deflector field of view. Input bus light is tapped by selected output fibers that are arranged at the appropriate positions on the I/O face of a GRIN lens. This version provides a set-and-forget feature; i.e., once the orientation of the stripes has been established by the field coils to direct input light to the selected tap, all coil current can be removed until another tap fiber selection is required.

Meaurements and observations to date indicate that the switch has a number of other desirable features. It is insensitive to incident polarization, making it particularly attractive for fiber optics. Crosstalk can be held to < - 20 db. with a fanout of 10. Physical size can be as low as 1 cm^3 as shown in figure 1, a working model of a 1 X 3 switch. Field coil switching is approximately 1 watt with microsecond select time. As a diffraction grating it satisfies the reciprocity theorem and is dispersive, allowing two-way operation of multiwavelength carriers. These attributes provide significant systems potential.

One useful systems device is an optical switch that can access just one output or simultaneously access all outputs, on electrical command. It was demonstrated that the basic stripe domain fiber optic switch can accomplish this because of its adaptive optic capability. With reference to figure 4b, if the stripe grating is switched to the constant periodicity radial format discussed under adaptive optics, the deflection space is a thin annulus at the same deflection angle as for the linear grating. Since the tap fibers are arranged to intercept light at this angle, they are uniformly illuminated with the light propagating in the bus fiber. Thus, all tap fibers are simultaneously selected. This can be implemented by placing a point magnetic pole on the optic axis, just behind the mirror. Switching between single select and the multiselect star coupler mode should occur in microseconds.

Because of its high fanout the garnet fiber optic switch may have use as a residue arithmetic adder for optical computing. Figure 8 is a schematic of a possible modulo 5 adder in which input fiber light represents one addend and the grating selected output represents the other.

SUMMARY

Magnetic stripe domain arrays in bismuth iron garnet epitaxial films are alterable in orientation and grating constant with externally applied magnetic fields generated by high speed deflection circuitry. Optical radiation, incident upon the grating, is thus diffracted in a two-dimensional solid-state fashion. Favorable deflection efficiencies and local control of the grating suggest a number of applications. Further work may lead to compact devices that are applicable to robotics and optical processing.

REFERENCES

1. Johansen, T. R.; Norman, D. I.; Torok, E. J.: Variation of Stripe Domain Spacing in a Faraday Effect Light Deflector. JAP, Vol. 42, No. 4, Mar. 1971, 1715.
2. Scott, G. B.; Lacklison, D. E.: Magneto-optic Properties and Applications of Bismuth Substituted Iron Garnets. IEEE Trans. on Mag., Vol. Mag-12, No. 4, July 1976, 292.
3. Fleming, D. L.; Lund, R. E.: Magneto-optic Bubble Domain Spatial Light Modulator SPIE, Vol. 128, 1977, 230.
4. Dimmock, J. O.; Huorwitz, C. E.; Reed, T. B.: Infrared Transmission Magnetic Birefringence and Faraday Rotation in EuO. APL, Vol. 14, No. 2, 15 Jan., 1969.
5. Ahn, K. Y.; Suits, J. C.: IEEE Trans. on Mag., Vol. MAG-3, Sept., 1967.
6. Shafer, A. W.; Torrance, J. B.; Penney, T.: Correlations of Infrared Absorption, Conductivity, and Stoichiometry in EuO. Magnetism and Magnetic Materials, C. D. Graham, Jr. (ed.), AIP, 1971.
7. Soref, R. A.; Kestigian, M.: Development of Liquid Crystal and Magnetic Stripe-Domain Multimode Optical Switches. Rome Air Development Center, Report SRC-CR-80-25, 1980.
8. Nelson, G. L.; Harvey, W. A.: Optical Absorption Reduction in $Bi_1Lu_2Fe_5O_{12}$ Garnet Magneto-Optical Materials, JAP, February 1982. (To be published)

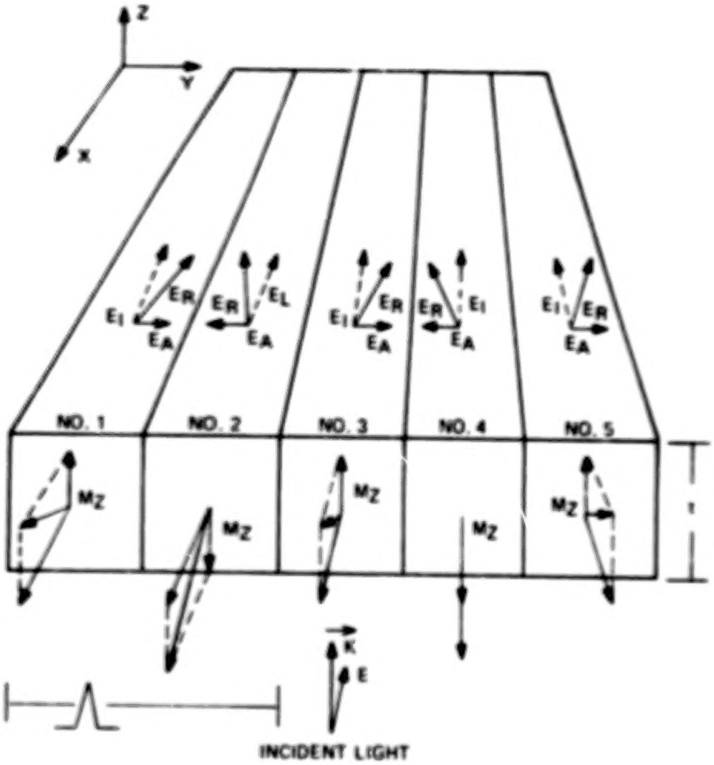

Z
Y
X
E_I
E_R
E_A
E_R
E_L
E_A
E_I
E_R
E_A
E_R
E_I
E_A
E_I
E_R
E_A
NO. 1
NO. 2
NO. 3
NO. 4
NO. 5
M_Z
M_Z
M_Z
M_Z
M_Z
t
K
E
INCIDENT LIGHT

Figure 1.- Magneto-optic diffraction grating.

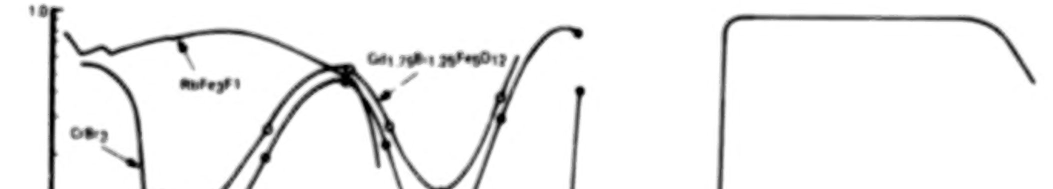


Figure 2.- Calculated diffraction efficiency of stripe domain materials.

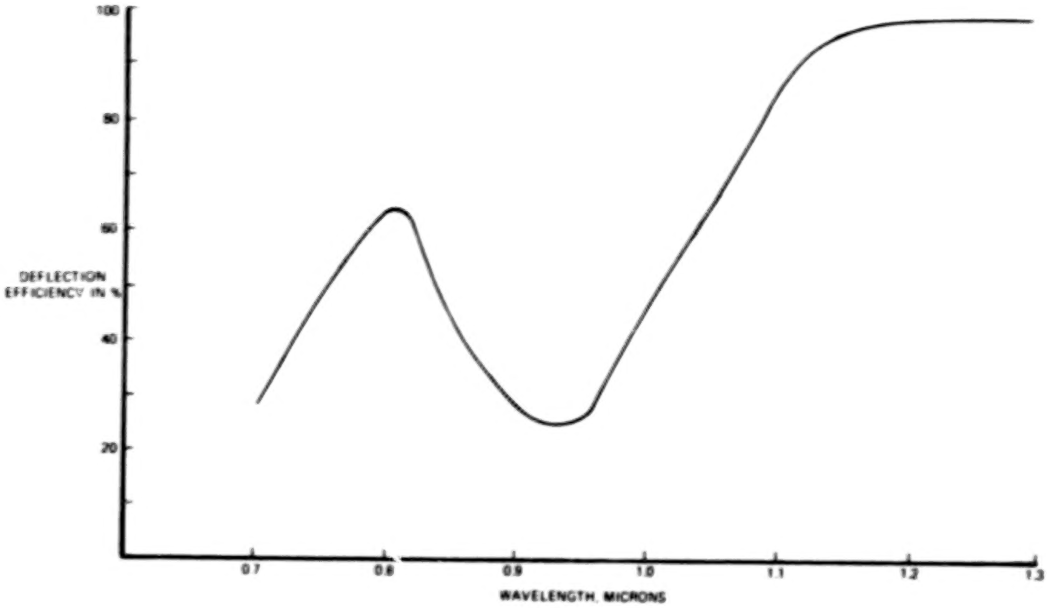


Figure 3.- Calculated diffraction of optimized garnet film.

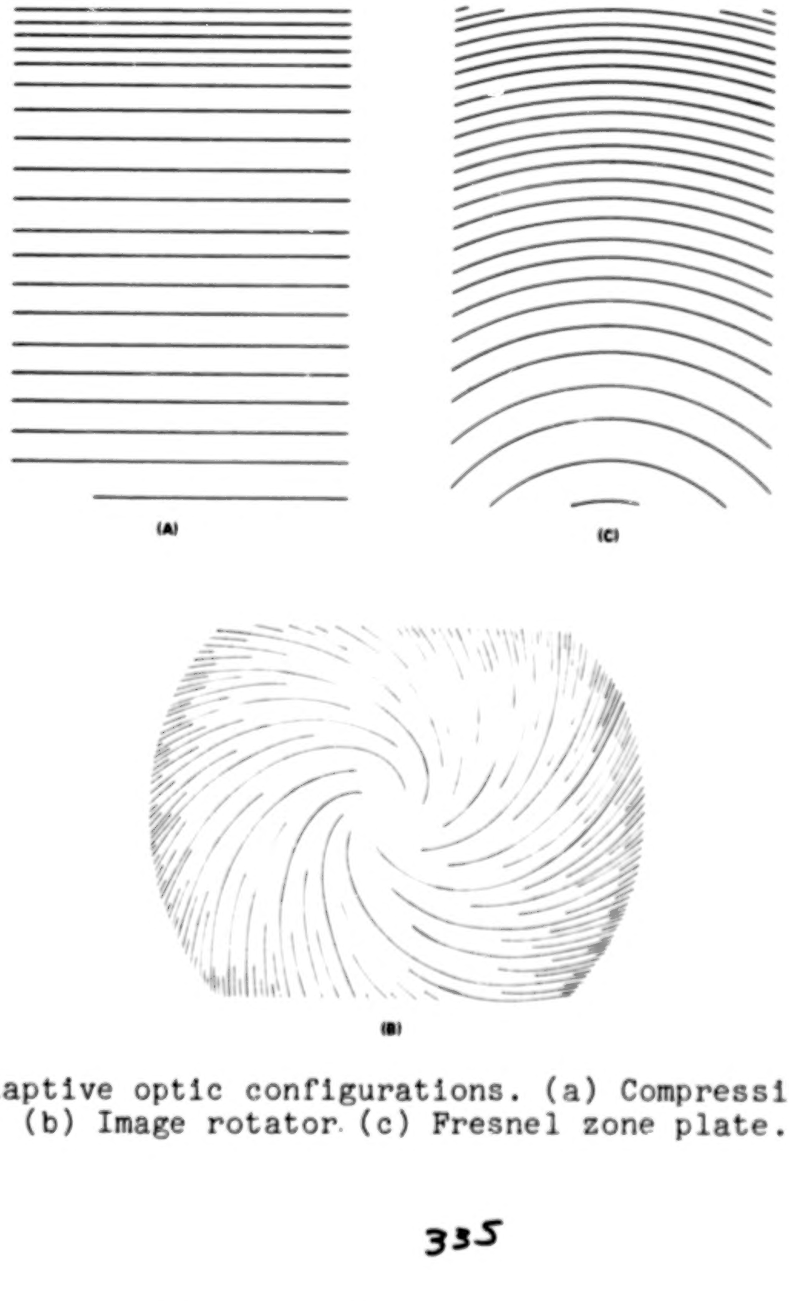


Figure 4.- Adaptive optic configurations. (a) Compression/expansion. (b) Image rotator. (c) Fresnel zone plate.

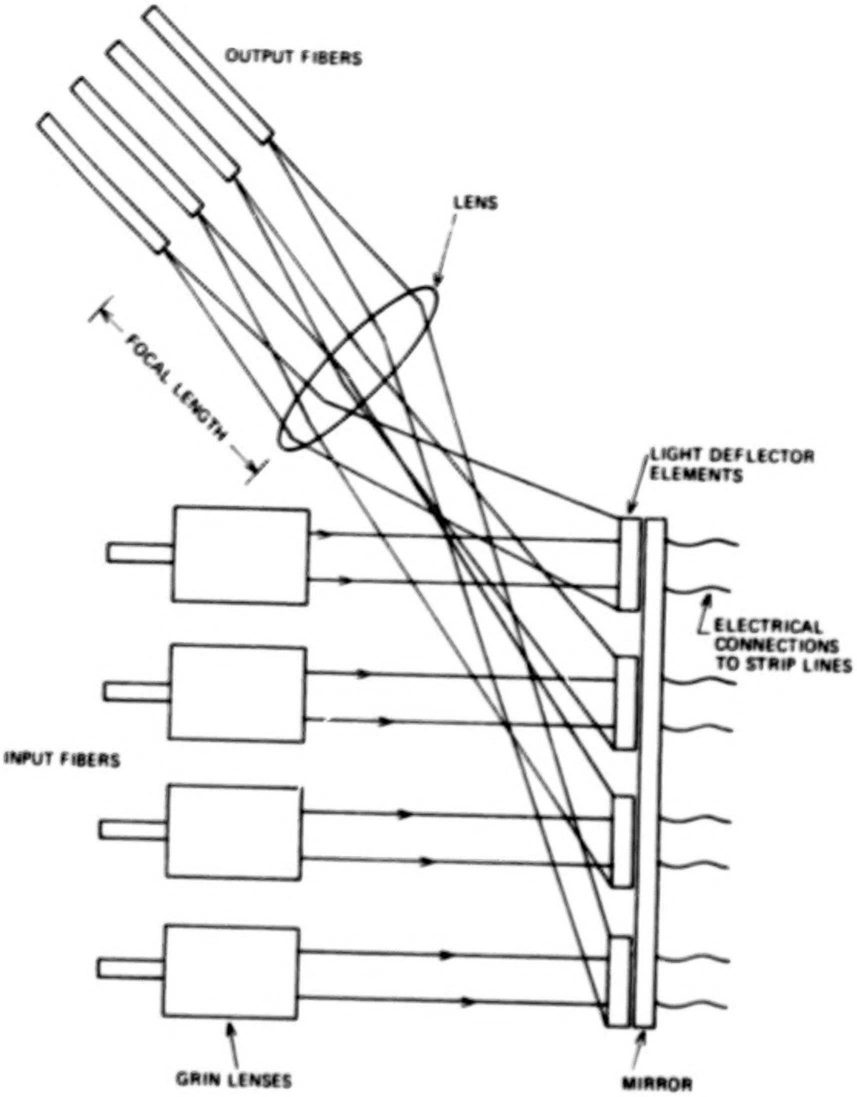

OUTPUT FIBERS
LENS
FOCAL LENGTH
LIGHT DEFLECTOR ELEMENTS
ELECTRICAL CONNECTIONS TO STRIP LINES
INPUT FIBERS
GRIN LENSES
MIRROR

Figure 5.- N X M swithboard.

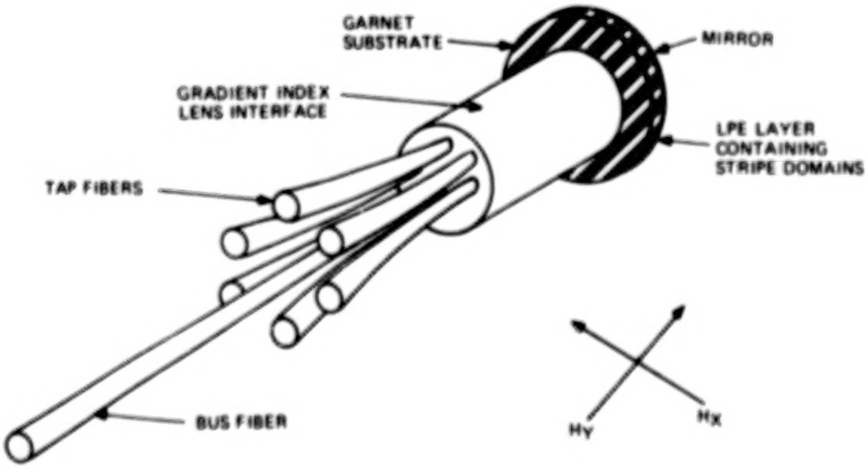

GARNET SUBSTRATE
MIRROR
GRADIENT INDEX LENS INTERFACE
LPE LAYER CONTAINING STRIPE DOMAINS
TAP FIBERS
BUS FIBER
H_Y
H_X

Figure 6.- Fiber optic switch.

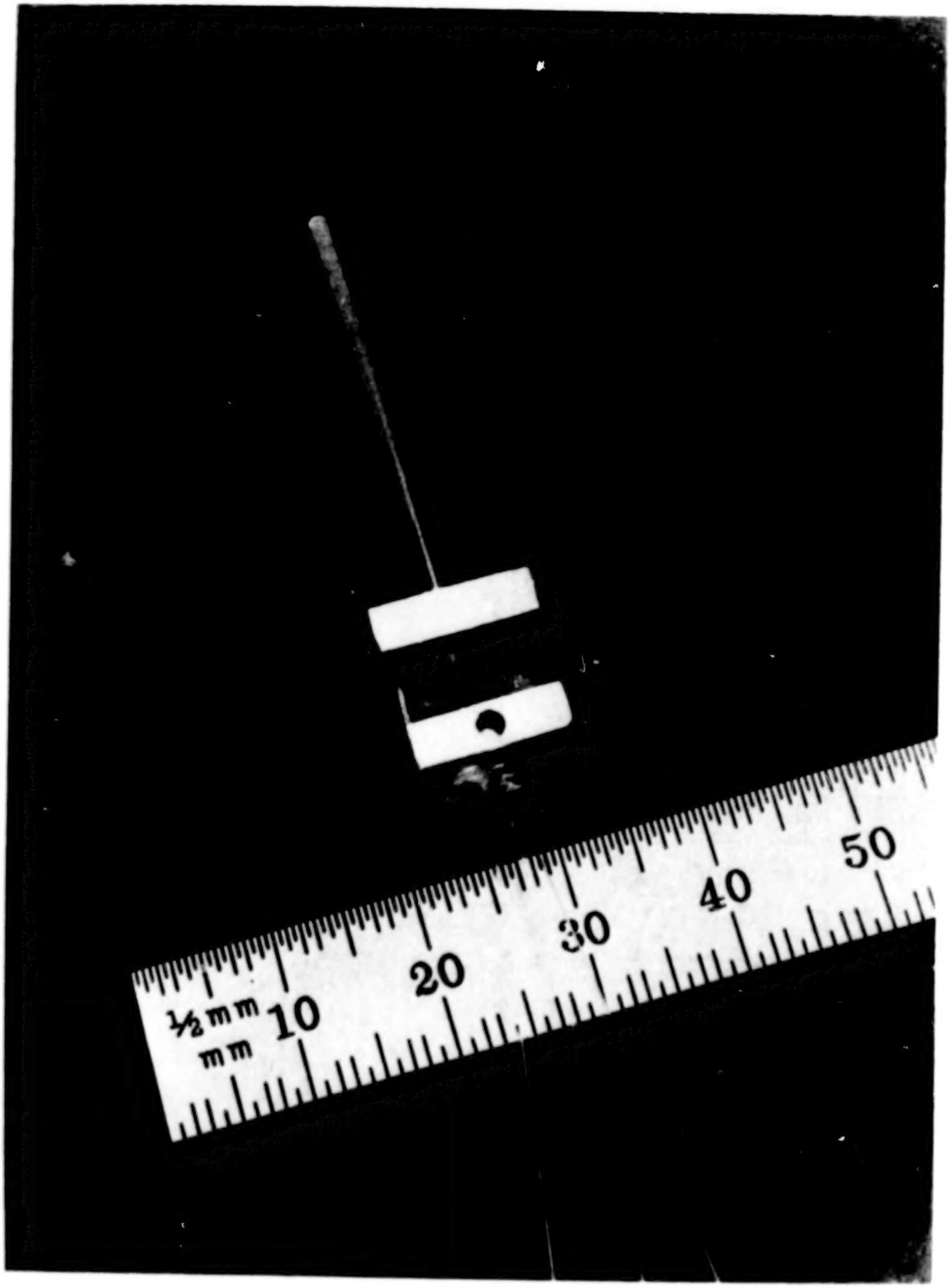

½mm
mm
10
20
30
40
50

Figure 7.- Operational model of 1 x 3 garnet fiber optic switch.

BEST DOCUMENT AVAILABLE

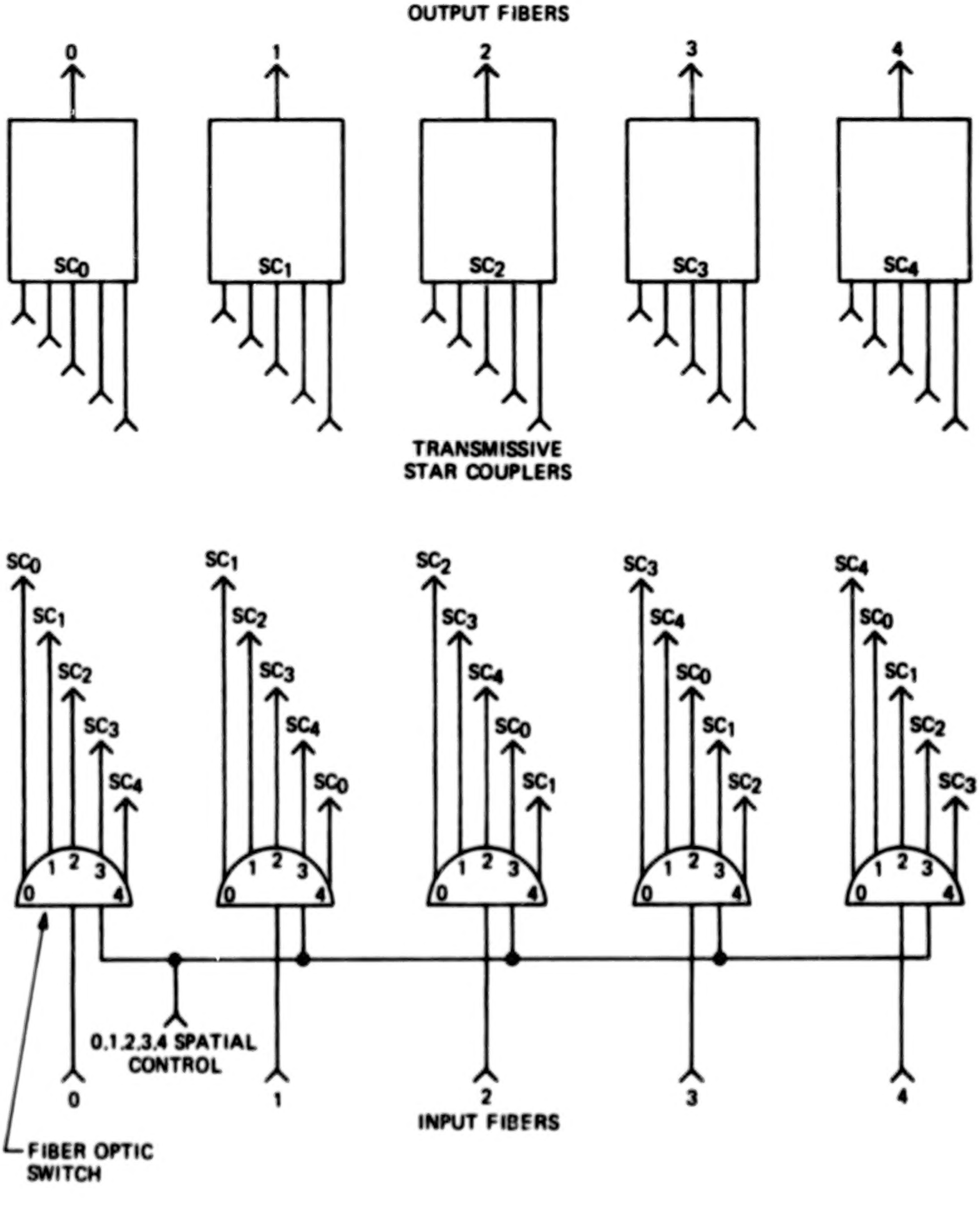

OUTPUT FIBERS
0
1
2
3
4
SC_0
SC_1
SC_2
SC_3
SC_4
TRANSMISSIVE STAR COUPLERS
SC_0 SC_1 SC_2 SC_3 SC_4
SC_1 SC_2 SC_3 SC_4 SC_0
SC_2 SC_3 SC_4 SC_0 SC_1
SC_3 SC_4 SC_0 SC_1 SC_2
SC_4 SC_0 SC_1 SC_2 SC_3
0 1 2 3 4
0,1,2,3,4 SPATIAL CONTROL
0
1
2
3
4
INPUT FIBERS
FIBER OPTIC SWITCH

Figure 8.- Residue arithmetic modulo 5 adder.

EFFICIENT BUTT COUPLING OF HIGH POWER CDH-LOC LASERS TO Ti-LiNbO$_3$ (LNT) OPTICAL WAVEGUIDES*

J. M. Hammer, D. Botez, C. C. Neil and J. C. Connolly
RCA Laboratories
David Sarnoff Research Center
Princeton, NJ 08540

We have observed direct butt coupling of over 50 percent of the light from CDH-LOC lasers into both Ti indiffused LiNbO$_3$ optical waveguides. This is the highest direct coupling of junction laser light to the technologically important class of indiffused LiNbO$_3$ type waveguides. Such waveguides are in use for a variety of optical switching and processing applications. In particular, compact high-frequency spectrum analyzers using Ti-LiNbO$_3$ are in an advanced state of development for military applications. The RCA CDH-LOC lasers used in this work have been brought to a high degree of effectiveness. These lasers provide single spatial mode output at the highest CW power (40 mW) reported.

Efficient butt coupling requires that the beam waist size of the laser be matched to that of the optical waveguide. The CDH-LOC laser is ideal for this purpose in that large beam waists commensurate with those of the waveguides are available. Thus, by producing lasers with the appropriate beam waist size the high coupling efficiency reported here is obtained.

The experiment consists of measuring the light coupled out of the waveguide by a prism coupler, when laser light is coupled into the W.G. The coupling is accomplished as follows: the laser junction plane is aligned with the waveguide plane and the laser output facet brought into close proximity to the polished waveguide edge using a high resolution-six-degree-freedom positioner.

---

* Because this paper was not available at the time of publication, only the abstract is included.

Both the experimental arrangement of the coupler and our observations will be described in detail. The outstanding features of the CDH-LOC lasers will also be reviewed.

This work is partially supported by NASA and NRL.

ATTENDEES

R. Martin Aherron
Information and Control Systems, Inc.
28 Research Drive
Hampton, VA 23666
(804) 838-9371

Frank Allario
NASA Langley Research Center
Mail Stop 476
Hampton, VA 23665
(804) 827-3601

Robert F. Arduini
Computer Sciences Corporation
Hampton, VA 23666
(804) 826-1725

Ravindra A. Athale
Naval Research Lab
Code 6530
Washington, DC 20375
(202) 767-5630

Joseph M. Ballantyne
Cornell University
214 Phillips Hall
Ithaca, NY 14853
(607) 256-4109

Michael K. Barnoski
TRW Technology Research Center
2525 E. El Segundo
El Segundo, CA 90245
(213) 535-4198

Ted E. Batchman
University of Virginia
Electrical Engineering Department
Charlottesville, VA 22901
(804) 924-3279

Robert J. Baumbick
NASA Lewis Research Center
21000 Brookpark (MS 100-1)
Cleveland, OH 44135
(216) 433-4000, Ext. 6136

Marvin E. Beatty III
NASA Langley Research Center
Mail Stop 473
Hampton, VA 23665
(804) 827-3535

W. Donald Bennettson
Grumman Aerospace Corporation
S. Oyster Bay Road
Bethpage, NY 11714
(516) 575-3777

Norman J. Berg
2800 Powder Mill Road
Adelphi, MD 20783
(202) 394-2520

Thomas J. Bicknell
Jet Propulsion Laboratory
4800 Oak Grove Drive
Pasadena, CA 91103
(213) 354-2523

Arthur E. Bisson
TRW
7600 Colshire Drive
McLean, VA 22102
(703) 734-6150

Robert W. Brandstetter
Grumman Aerospace
Bethpage, NY 11714
(516) 575-3528

Robert H. Braunstein
CIA
Washington, DC 20505
No telephone number listed

Alpheus W. Burner
NASA Langley Research Center
Mail Stop 236
Hampton, VA 23665
(804) 827-3234

James R. Busch
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201
(614) 424-4381

Richard F. Carson
Texas Tech University
Box 4439
Lubbock, TX 79409
(806) 742-3532

David P. Casasent
Carnegie-Mellon University
Pittsburgh, PA 15213
(412) 578-2464

Steven K. Case
University of Minnesota
123 Church Street
Minneapolis, MN 55455
(612) 373-5627

H. John Caulfield
Innovative Optics, Inc.
P. O. Box 1275
Concord, MA 01742
(617) 263-0482

Donald J. Channin
RCA Corporation
Princeton, NJ 08540
(609) 734-2383

Jar-Mo Chen
Martin Marietta Laboratories
1450 S. Rolling Road
Baltimore, MD 21227
(301) 247-0700

Gordon Chin
NASA Goddard Space Flight Center
Code 693
Greenblet, MD 20771
(301) 344-5333

Chuck R. Chubb
McDonnell Douglas Astronautics Co.
St. Louis, MO
(314) 232-6116

Dewey L. Clemmons, Jr.
NASA Langley Research Center
Mail Stop 249A
Hampton, VA 23665
(804) 827-3216

Dennis S. Cleverly
General Electric Company
French Road (MD 417)
Utica, NY 13501
(315) 797-1000

James J. Connors
Perkin-Elmer Corporation
Mail Stop 884
Danbury, CT 06810
(203) 797-6389

J. Thomas Cox
NV&EOL
Ft. Belvoir, VA 22060
(703) 664-5468

Roger K. Crouch
NASA Langley Research Center
Hampton, VA 23665
(804) 827-3535

Q. Isa Daudpota
NASA Langley Research Center
Mail Stop 479
Hampton, VA 23665
(804) 827-3029

Dale R. Davidson
Sperry Univac
1745 S. Jefferson Davis Highway
Arlington, VA 22202
No telephone number listed

Arthur J. Decker
NASA Lewis Research Center
Mail Code 77-1
21000 Brookpark Road
Cleveland, OH 44135
(216) 433-4000, Ext. 353

Antonio R. Dias
Environmental Research Institute of Michigan
P. O. Box 8618
Ann Arbor, MI 48107
(313) 994-1200

A. Robert Doucette
Grumman Aerospace
C20-25
Bethpage, NY 11714
(516) 575-6128

Frederick L. Drain
NASA Langley Research Center
Mail Stop 477
Hampton, VA 23665
(804) 827-3681

Sverre T. Eng
Chalmers University of Technology
S-41296 Göteborg, Sweden
(031) 81 01 00

Warren O. Essler
Tennessee Tech. University
Cookeville, TN 38501
(615) 528-3460

Dick Fikes
Naval Electronic System Command
National Center 1
Washington, DC 20360
(202) 692-6412

Arthur G. Foyt
MIT Lincoln Laboratories
244 Wood Street
Lexington, MA 02173
(617) 862-5500

John M. Franke
NASA Langley Research Center
Mail Stop 236
Hampton, VA 23665
(804) 827-3234

Randy W. Frey
Martin Marietta
Sand Lake Road
Orlando, FL 32855
(305) 352-4499/851-4242

Frederick W. Freyre
Hazeltine Corporation
Cuba Hill Road
Greenlawn, NY 11740
(516) 265-7000

Archibald L. Fripp
NASA Langley Research Center
Mail Stop 473
Hampton, VA 23665
(804) 827-3535

F. Trevor Gamble
Denison University
Granville, OH 43023
No telephone number listed

John H. Goebel
NASA Ames Research Center
Moffett Field, CA 94035
(415) 965-6525

Nicholas S. Gorton
Computer Sciences Corporation
Hampton, VA 23666
(804) 826-1725

Ronald Gorton
Kentron International/NASA Langley
Mail Stop 470
Hampton, VA 23665
(804) 827-3535

William R. Graver
Riverside Research Institute
1701 N. Ft. Myer Drive
Arlington, VA 22209
(703) 522-2310

Alex Grumet
Grumman Aerospace Corporation
Mail Stop A08-35
Bethpage, NY 11714
(516) 575-2530

William Guyton
Lockheed Electronics
1501 U. S. Highway 22
Plainfield, NJ 07061
(201) 757-1600

Maurice Halioua
New York Institute of Technology
Old Westbury, NY 11568
(516) 686-7733

Nesim Halyo
Information and Control Systems, Inc.
28 Research Drive
Hampton, VA 23666
(804) 838-9371

Jacob M. Hammer
RCA Laboratories
Princeton, NJ 08540
(609) 734-3210

Michael M. Hart
NASA Langley Research Center
Mail Stop 471
Hampton, VA 23665
(804) 827-3535

Paul J. Hayes
NASA Langley Research Center
Mail Stop 470
Hampton, VA 23665
(804) 827-3535

Dessie M. Hendricks
7 W. River Road
Poquoson, VA 23662
(804) 827-3418

Herbert D. Hendricks
NASA Langley Research Center
Mail Stop 499
Hampton, VA 23665
(804) 827-3418

William C. Honaker
NASA Langley Research Center
Mail Stop 479
Hampton, VA 23665
(804) 827-3029

Alan Huang
Bell Laboratories
Room 4C
520 Crawfords Corner Road
Holmdel, NJ 07733
(201) 949-7631

Friedrich O. Huck
NASA Langley Research Center
Mail Stop 473
Hampton, VA 23665
(804) 827-3535

Lynn D. Hutcheson
Honeywell Corporate Technology Center
10701 Lyndale Avenue S.
Bloomington, MN 55420
(612) 887-4052

Guy Indebetouw, Jr.
VPI & SU
Blacksburg, VA 24060
(703) 951-5767

Howard A. Jenkinson
U. S. Army ARRADCOM
DRDAR-SCF-10
Dover, NJ 07801
(201) 328-2936

Daniel J. Jobson
NASA Langley Research Center
Mail Stop 471
Hampton, VA 23665
(804) 827-3535

David R. Johnson
NASA Langley Research Center
Mail Stop 235
Hampton, VA 23665
(804) 827-3907

James W. Johnson
NASA Langley Research Center
Mail Stop 490
Hampton, VA 23665
(804) 827-3631

Alan R. Johnston
Jet Propulsion Laboratory
Mail Stop 189-326
4800 Oak Grove Drive
Pasadena, CA 91103
(213) 354-4054

Robert L. Kenimer
10 Belmont Place
Hampton, VA 23666
(804) 826-0394

Stephen T. Kowel
Syracuse University
111 Link Hall
Syracuse, NJ 13210
(315) 423-4453

Peter R. Kurzhals
NASA Goddard Space Flight Center
Code 500
Greenbelt, MD 20771
(301) 344-6521

John N. Lee
Naval Research Lab
Code 6530
Washington, DC 20375
(202) 767-3100

Sing H. Lee
University of California at San Diego
La Jolla, CA 92093
(714) 452-2413

Craig M. Lemrow
Corning Glass Works
Baron Steuben-6
Corning, NY 14831
No telephone number listed

William D. Mace
NASA Langley Research Center
Mail Stop 117
Hampton, VA 23665
(804) 827-3745

Carl J. Magee
NASA Langley Research Center
Mail Stop 499
Hampton, VA 23665
(804) 827-3418

Brian James Markey
COMSAT Labs
22300 Comsat Drive
Clarksburg, MD 20871
(301) 428-4400

Duncan E. McIver, Jr.
NASA Langley Research Center
Mail Stop 476
Hampton, VA 23665
(804) 827-3601

Glen M. McWright
University of Virginia
Dept. of Electrical Engineering
Charlottesville, VA 22901
(804) 924-3279

Susan L. Meardon
WR-ALC/MMECE
Robins AFB
Warner Robins, GA 31098
AV 468-4669

Sheldon Meth
The BDM Corporation
7915 Jones Branch Drive
McLean, VA 22102
(703) 821-5163

William E. Miller
NASA Langley Research Center
Mail Stop 499
Hampton, VA 23665
(804) 827-3418

William M. Moore
NASA Langley Research Center
Mail Stop 476
Hampton, VA 23665
(804) 827-3601

John A. Neff
AFOSR/NE
Bolling AFB
Washington, DC 20332
(202) 767-4933

Charles D. Nichols
NASA Langley Research Center
Mail Stop 470
Hampton, VA 23665
(804) 827-3535

Stewart L. Ocheltree
NASA Langley Research Center
Mail Stop 235A
Hampton, VA 23665
(804)827-2791

José A. Ortiz
NASA Langley Research Center
Mail Stop 470
Hampton, VA 23665
(804) 827-3535

Thomas C. Payne
Wyle Laboratories
3200 Magruder Boulevard
Hampton, VA 23666
(804) 838-5000

Melvin E. Pedinoff
Hughes Research Labs
3011 Malibu Canyon Road
Malibu, CA 90265
456-6411

Demetri Psaltis
California Institute of Technology
Mail Code 116-81
Pasadena, CA 91125
(213) 356-4856

Joseph J. Rebholz
Honeywell, Inc.
10701 Lyndale Avenue S.
Bloomington, MN 55420
(612) 887-4026

William R. Reed
NASA Johnson Spacecraft Center
Mail Stop EH4
Houston, TX 77058
(713) 483-2801

Donald Sashin
University of Pittsburgh
RC 406 Scaife Hall
Pittsburgh, PA 15261
(412) 647-3490

Cesar A. Sepulveda
Jet Propulsion Laboratory
Mail Stop 169-315
4800 Oak Grove Drive
Pasadena, CA 91103
(213) 354-4612

Wai-Pong Shing
VPI & SU
Blacksburg, VA 24060
(703) 961-7054

Jag J. Singh
NASA Langley Research Center
Mail Stop 235
Hampton, VA 23665
(804) 827-3907

G. Louis Smith
NASA Langley Research Center
Mail Stop 401B
Hampton, VA 23665
(804) 827-2868

Martin M. Sokoloski
NASA Headquarters
Mail Code RTE-6
Washington, DC 20546
(202) 755-8505

Joseph D. Spear-Zino
Perkin-Elmer Corporation
100 Wooster Heights Road
Mail Stop 813
Danbury, CT 06810
(203) 797-5639

J. Larry Spencer
NASA Langley Research Center
Mail Stop 477
Hampton, VA 23665
(804) 827-3681

Robert B. Spiers, Jr.
NASA Langley Research Center
Mail Stop 473
Hampton, VA 23665
(804) 827-3535

Carl Spight
AMAF Industries, Inc.
P. O. Box 1100
Columbia, MD 21044
No telephone number listed

Robert L. Stermer, Jr.
NASA Langley Research Center
Mail Stop 470
Hampton, VA 23665
(804) 827-3535

Harold H. Szu
Naval Research Lab
Washington, DC 20375
(202) 767-2407

Armand R. Tanguay, Jr.
University of Southern California
523 Seaver Science Center
Los Angeles, CA 91104
(213) 743-6152

Andrew Tarasevich
Lockheed Electronics
1501 U. S. 22
Plainfield, NJ 07061
(201) 757-1600

Ernest J. Torok
Univac
Box 3525
St. Paul, MN 55165
(612) 456-2397

Wilfrid B. Veldkamp
MIT Lincoln Laboratory
Box 73
Lexington, MA 02173
(617) 862-5500, Ext. 7579

Carl M. Verber
Battelle Columbus Laboratories
505 King Avenue
Columbus, OH 43201
(614) 424-5710

Cardinal Warde
MIT
Room 13-3134
Cambridge, MA 02139
(617) 253-6858

John E. Weaver
Marconi Avionics, Inc.
4500 North Shallowford Road
Atlanta, GA 30338
(404) 394-7800, Ext. 285

David E. Weissman
Hofstra University
Hempstead, NY 11550
No telephone number listed

Robert G. Wilson
NASA Langley Research Center
Mail Stop 494
Hampton, VA 23665
(804) 827-3856

Neil Zepkin
Lockheed Electronics Company
1501 U. S. Rt. 22, MS 239
Plainfield, NJ 07061
(201) 757-1600

1. Report No. NASA CP-2207	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle OPTICAL INFORMATION PROCESSING FOR AEROSPACE APPLICATIONS		5. Report Date December 1981
		6. Performing Organization Code 506-54-63-03
7. Author(s)		8. Performing Organization Report No. L-15018
		10. Work Unit No.
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665		11. Contract or Grant No.
		13. Type of Report and Period Covered Conference Publication
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The proceedings of the Conference on Optical Information Processing for Aerospace Applications held on August 18-19, 1981, in Hampton, Virginia, are presented. Initial papers indicate potential applications of optical processing in the areas of pattern recognition, aircraft and spacecraft control, and robotics. Detailed component and device papers review the state of optical circuit technology. Matrix mathematics are shown to be ideally suited for implementation in optical hardware. Optical information processing offers significant potential for meeting processing requirements of projected aerospace missions.

17. Key Words (Suggested by Author(s)) Integrated optics Pattern recognition Optical computers Optical components Optical devices	18. Distribution Statement Unclassified - Unlimited Subject Category 74

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 365	22. Price A16

For sale by the National Technical Information Service, Springfield, Virginia 22161

348